

**Isolering og karakterisering
av et operon for chaperoner fra bakterien
Chloroflexus aurantiacus
av
Elisabeth Ytteborg**

**Hovedfagsoppgave i generell mikrobiologi for graden Candidata scientiarum ved
Institutt for molekylær biovitenskap, Universitetet i Oslo**

2005

FORORD

Arbeidet i denne oppgaven ble utført ved Institutt for molekylær biovitenskap, Universitetet i Oslo, i perioden høsten 2003 til våren 2005.

Lik *Chloroflexus aurantiacus*, den lille bakterien denne oppgaven omhandler, som fant sin nisje i de varme kildene på Island, planeten jorden, må vi mennesker også finne vår lille nisje på den samme planeten. Etter jordens 6 milliarder års utvikling og artenes 3,5 milliarder år, har menneskene skapt seg en tilværelse, et samfunn, der valg av nisje verken er enkelt eller rettferdig. Er man av de heldige i dette samfunnet, har man muligheten til selv å velge hvor man vil havne. Basert på interesser, evner, ferdigheter, miljø og en stor grad tilfeldigheter ender vi alle opp et sted. Blandingen av dette førte meg til rom 3602 i 3de etasje i Kristine Bonnevis hus, Universitetet i Oslo, Blindern. Jeg retter en stor takk til min veileder, professor Reidun Sirevåg, som på grunn av sin interesse ga meg oppgaven om den lille varmeresistente bakterien. Oppgaven har gledet meg og gitt meg kunnskaper, ikke bare innenfor mikrobiologi og molekylærbiologi, men også lærdom om oss mennesker og vårt evige drag til å vite. Jeg vil også takke professor William Davies, tidligere stipendiat Alexandra Björk, stipendiat Marc Anglès d'Auriac og mine medstudenter på rom 3602.

Til slutt vil jeg takke organismen *Chloroflexus aurantiacus* og alle de millionene av celler som ble ofret for at denne oppgaven skulle kunne ta form.

Takk

Oslo, juli 2005

Elisabeth Ytteborg

SAMMENDRAG

I denne oppgaven var målet å isolere genet som koder for chaperoninet GroESL1 i *C. aurantiacus* fra et genomisk bibliotek av bakterien. Genet som koder for GroESL2 var tidligere isolert fra det samme biblioteket (Vanberg, Hovedfagsoppgave UiO, 2001) ved hjelp av PCR. Denne metoden ble derfor valgt.

Da oppgaven ble påbegynt var ikke genomet til *C. aurantiacus* ferdigsekvansert, men en stor del av genomet lå ute på NCBI som "shot-gun" sekvenser. Imidlertid var noe av sekvensen til *groESL1* allerede kjent, Chlo_724 (figur 4.20) var på 1430 bp. Da promoterområdet til operonet og *groES1* utgjorde 582 av disse basene, og forventet størrelse til *groEL1* var ca. 2000 bp, manglet sekvensen ca. 1000 bp.

Spesifikke primere for genet ble utformet og en PCR basert søkemethode etter King (1997) sammen med en noe modifisert metode etter Yu og Bloem (1996) ble brukt. Med disse metodene ble flere positive fagkloner isolert fra biblioteket. Innskuddene fra de positive fagklonene ble videre analysert med PCR med kombinasjonen av vektorspesifikke primere og genspesifikke primere for å finne ut om hele *groESL1* var tilstede i innskuddet.

Sekvenseringen av innskuddene ble gjort av firmaene GATC og MWG. Flere innskudd ble isolert og sekvensert, men den fullstendige sekvensen til *groESL1* ble ikke funnet. En klon, kalt Plakk2, ga en fortsettelse av sekvensen Chlo_724 på ca. 800 bp. Denne sekvensen ble analysert i programmet WebCutter. I dette programmet ble sekvensen kuttet *in silico* med restriksjonsenzymet *Sau3AI*, som var enzymet som var brukt til å kutte genomisk DNA fra *C. aurantiacus* da biblioteket ble konstruert. Resultatet av analysen viste 17 kuttsteder i området der sekvensene sluttet.

Høsten 2004 ble genomet til *C. aurantiacus* ferdig sekvensert og lagt ut på databasene til NCBI som en "shot gun" sekvens i form av "contigs", delsekvenser. Siste del av Plakk2, omtrent 200 bp, ble brukt i BLAST søk. En klon, Chlo01_574 viste 100 % sekvenslikhet med den siste delen av Plakk2. Sekvensen som ligger mellom det som var kjent fra før og den siste delen av Plakk2 var imidlertid ikke å finne i databasen. Muligens var det her et glipp i sekvenseringen fra TIGR. Sekvensen ble satt sammen så de utgjorde hele operonet *groESL1* (figur 4.24). Sekvensen til genet var på 2228 bp og inneholdt stopp-kodonet TAA. Sekvensen ble oversatt til proteinsekvens og molekylvekten beregnet ved hjelp av programmet ExPacy. Molekylvekten på dette proteinet var 59,7 kDa.

Det ble utformet primere for å isolere hver av de to genene i operonet, *groES1* og *groEL1*, ved hjelp av PCR. Primersettene ble brukt både med kromosomalt DNA fra og biblioteket som templat. Primersettet til *groES1* ga gode resultater med begge templatene, primersettet til *groEL1* ga kun resultat med kromosomalt DNA som templat. Dette styrket teorien om at hele sekvensen til *groEL1* ikke var å finne i en klon. PCR produktene ble sekvensert og viste seg å stemme med sekvensene i operonet.

De to genene skulle uttrykkes og ble derfor satt inn i hver sin pET vektor som ble brukt til å transformere BL21 Star celler. Det viste seg at det bare var *groES1* som ble uttrykt på denne måten. Sekvensen til *groEL1* ble videre analysert i programmet Rare Codon Calculator der det ble funnet at dette genet hadde i alt 14 sjeldne kodoner for aminosyre prolin og arginin. Sjeldne kodoner kan være årsaken til at genet ikke blir uttrykt, eller at det blir for tidlig terminering av heterologe gener. Det ble derfor forsøkt å uttrykke *groEL1* i vertscellen BL21-CodonPlus (DE3) RP. Stammen BL21-CodonPlus (DE3) RP inneholder ekstra kopier av *argU* og *proL*, genene som koder for tRNA som

gjenkjenner kodonene AGA og AGG for arginin og kodonet CCC for prolin. Da IPTG induisert celleekstrakt fra disse cellene ble analysert på SDS-PAGE, viste det seg at et protein på 60 kDa var uttrykt.

INNHALDSFORTEGNELSE

Forord	2
Sammendrag	3
Bakgrunn for oppgaven	8
Mål og hensikt med oppgaven	9
Forkortelser	10
1. Innledning	11
1.1 Livets stamtre	11
1.1.1 <i>Chloroflexus aurantiacus</i>	12
1.2 Termofile organismer	14
1.2.1 Termostabilitet	15
1.2.2 Varmesjokkrespons	17
1.2.3 Molekylære chaperoner	20
1.2.4 Chaperoniner	21
1.3 Isolering av gener ved bruk av PCR	26
1.3.1 Sekvensering og sekundære strukturer i DNA	27
2. Materialer	31
2.1 Kjemikalier	31
2.1.1 Uorganiske kjemikalier	31
2.1.2 Organiske kjemikalier	31
2.2 Primere	32
2.3 Standarder	33
2.4 Enzymer	33
2.5 Utstyr	33
2.6 Bakteriestammer og bakteriofag	34
2.7 Ferdige sett	34
2.8 Dataprogrammer	34
2.9 Dyrkningsmedier	35
2.9.1 Dyrkningsmedium for <i>C. aurantiacus</i>	35
2.9.2 Dyrkningsmedium for <i>E. coli</i>	36
2.9.3 Tilsetninger i dyrkningsmedier	36
2.10 Buffere og løsninger	37
2.10.1 Buffere for oppbevaring av fag	37
2.10.2 Løsninger for isolering av DNA	37
2.10.3 Buffere for oppbevaring av DNA	37
2.10.4 20 mM Kaliun-Fosfatbuffer pH 7.5	37
2.10.5 Buffere og løsninger for agarosegel og elektroforese	38
2.10.6 Løsninger for arbeid med proteiner	38
3. Metoder	40
3.1 Sterilteknikk	40
3.2 Dyrking av celler	40
3.2.1 Dyrking og høsting av <i>C. aurantiacus</i>	40
3.2.2 Dyrking av <i>E. coli</i>	40
3.2.3 Utsåing av bakteriofag λ ZAP	41

3.2.4 Lagring og oppbevaring av celler	41
3.2.5 Kompetente celler	42
3.3 Arbeid med DNA	42
3.3.1 Isolering av DNA fra hele celler	42
3.3.2 Isolering av kromosomalt DNA fra <i>C. aurantiacus</i>	43
3.3.3 Måling av DNA-konsentrasjon og undersøkelse av kvaliteten til isolert DNA	43
3.3.4 DNA til PCR reaksjon direkte på celler	44
3.3.5 Agarose gelelektroforese	44
3.3.6 Isolering av DNA fra agarosegel	45
3.4 Primerutforming og optimalisering av PCR	46
3.4.1 PCR	46
3.4.2 Primerutforming	47
3.4.3 Utforming av primere til søk etter positiv klon	47
3.4.4 Sekvenseringsprimer	47
3.4.5 Revers primer utformet fra <i>groESL2</i>	48
3.4.6 Konsensusprimer	48
3.5 Genbiblioteket	48
3.5.1 Konstruksjon av biblioteket	48
3.5.2 Test av biblioteket	49
3.5.3 Blå/hvit bakgrunnstest	49
3.5.4 Test av innskudd ved bruk av PCR	49
3.6 Søk i biblioteket ved bruk av PCR	50
3.6.1 Metoden til King	51
3.6.2 Metoden til Yu og Bloem	52
3.6.3 Isolering av innskudd fra positive fagklon	54
3.6.4 Test av innskudd fra positiv fagklon ved hjelp av PCR	55
3.7 Sekvensering av positive fagklon	56
3.8 Søk etter gen ved hjelp av T7 og genspesifikke forover primere	56
3.9 Isolering av <i>groES1</i> og <i>groEL1</i>	57
3.10 Transformerings og ekspressjon av <i>groES1</i> og <i>groEL1</i>	59
3.10.1 Prinsippet for retningsbesemt TOPO-kloning	59
3.10.2 Transformerings av <i>groES1</i> og <i>groEL1</i>	61
3.10.3 Ekspresjon av <i>groES1</i> og <i>groEL1</i>	62
3.11 SDS-PAGE	63
3.12 Sekvensering av <i>groEL1</i>	64
3.13 Databehandling av sekvenseringsresultatene	65
4. Resultater	67
4.1 Dyrking og høsting av <i>C. aurantiacus</i>	67
4.1.1 Isolering av kromosomalt DNA, måling av DNA-konsentrasjon og undersøkelse av kvaliteten til isolert DNA	67
4.2 Genbibliotek	68
4.2.1 Test av genbiblioteket	68
4.2.2 Optimalisering av primere til søk i bibliotek etter positiv fagklon	70
4.2.3 Søk etter positiv fagklon i genbiblioteket	71
4.2.4 Isolering av innskudd fra positiv fagklon	73
4.3 Søk etter gen ved hjelp av andre primer	75
4.3.1 Bruk av nye primere til "primerwalking"	76
4.3.2 Søk etter gen ved hjelp av T7 og genspesifikke forover primere	77
4.4 Analyse av sekvenseringsresultatene	77
4.5 Isolering av <i>groES1</i> og <i>groEL1</i>	82
4.6 Sekundære strukturer i DNA	84
4.7 Transformerings og ekspresjon av <i>groES1</i> og <i>groEL1</i>	86

5. Diskusjon	93
5.1 To operoner for <i>groESL</i>	93
5.2 Genbiblioteket	94
5.3 Sekvenseringsresultatene	97
5.4 Isolering, transformering og ekspresjon av <i>groES1</i> og <i>groEL1</i>	98
5.5 Sekundære strukturer i <i>groEL1</i>	100
5.6 Forslag til videre arbeid	101
Referanser	102
Appendiks	107

BAKGRUNN FOR OPPGAVEN

Chaperoninene GroEL (Hsp60) og GroES (Hsp10) er proteiner som hører til en undergruppe av molekylære chaperoner som er funnet i alle bakterier som er undersøkt. Chaperoninene er blant de mest konserverte proteinene vi kjenner og tilhører de såkalte varmesjokkproteinene. En av de rollene de har i cellene er å sørge for at polypeptider blir riktig foldet etter at de er syntetisert på ribosomene. En annen viktig rolle er å hindre aggregering og denaturering av proteiner når cellen er under stress som for eksempel høy temperatur. De to proteinene GroEL og GroES virker sammen og de tilsvarende genene *groEL* og *groES* er vanligvis organisert sammen i et operon, *groESL*, på bakteriekromosomet.

Den grønne, fototrofe, glidende bakterien *Chloroflexus aurantiacus* er en moderat termofil bakterie som har optimum- og maksimumtemperatur for vekst på henholdsvis 55 °C og 70 °C. Det er tidligere vist at når *C. aurantiacus* blir dyrket under forhold med økende temperatur, øker produksjonen av to polypeptider med molekylær vekt på 60 kDa og 10 kDa. Da disse polypeptidene ble isolert og den N-terminale enden sekvensert og analysert, viste de en høy grad av likhet med GroES og GroEL fra andre organismer. Sekvenseringen av den N-terminale enden av det rensede 60 kDa proteinet viste 55,6 % identisk identitet og 76 % likhet med GroEL fra *E. Coli*. For videre å karakterisere dette mulige GroEL-proteinet og dens korresponderende gen, ble et genomisk bibliotek med lambda- ZAP Express vektor laget. Dette biblioteket ble undersøkt ved hjelp av PCR, og et 600 bp fragment fra *C. aurantiacus* Hsp60 genet, som var isolert av andre forskere (Gupta et al.). Ved hjelp av denne metoden ble en positiv klon som inneholdt en åpen leserramme, tilsvarende et av operonene med de to chaperonin genene *groEL* og *groES*, isolert. Videre analyser av den N-terminale aminosyresekvensen, avledet fra den åpne leserrammen, viste at den ikke var identisk med den tidligere bestemte sekvensen fra det 60 kDa polypeptidet fordi 14 av de 45 aminosyrene var ulike. Hele genomet til *C. aurantiacus* var i ferd med å bli sekvensert og et BLAST søk viste at det var to operoner for chaperoniner i genomet til *C.aurantiacus*, et inneholder genene helt tilsvarende det tidligere isolerte proteinet, og et annet som var identisk med den åpne leserrammen isolert fra genbiblioteket. De to operonene er interessante ved at de har ulike oppstrøms reguleringsområder. I begge tilfellene er de to konsensus – 35 og – 10 sekvensene, som er funnet hos flere andre GroESL operoner, tilstede. Oppstrøms i det ene operonet, *groESL2*, er det i tillegg et såkalt CIRCE element. Videre ble det vist at *C. aurantiacus* under varmestress produserer store mengder av et protein med aminosyresekvens helt tilsvarende baserekkefølgen *groESL1*.

Tilstedeværelse av flere kopier av *groESL* gener kan ha betydning ved at de gir bakterien fleksibilitet ved varierende miljøbetingelser og fysiologiske krav. Genduplikering kan ha flere fordeler slik som økt ekspressjon av de gjeldende proteinene avhengig av hva slags forhold bakterien befinner seg i og en viss forsikring mot mutasjoner eller andre ødeleggende faktorer.

Et interessant spørsmål i forbindelse med dette er om de to operonene hos *C. aurantiacus* er regulert på ulike måter og om ekspresjonen skjer under forskjellige typer ytre stress forårsaket av temperatur, oksygen eller lys.

MÅL OG HENSIKT MED OPPGAVEN

Arbeidet som er presentert i denne oppgaven er en del av et større prosjekt rettet mot å undersøke det molekylære grunnlaget for termofili hos grønne fototrofe bakterier. En del av prosjektet går ut på å få mer informasjon om rollen til de to chaperoninene *GroESL1* og *GroESL2* hos bakterien *C. aurantiacus*. Et viktig spørsmål som ønskes besvart er om begge de to operonene er aktive, hva som regulerer om de kommer til uttrykk og i så fall om begge chaperoninene er funksjonelle. I denne oppgaven var målet å isolere operonet *groESL-1* ved å søke i *C. aurantiacus* genbiblioteket. Ved å isolere operonet kan genene uttrykkes i en ekspresjonsvektor. Videre studier kan dermed utføres. *GroESL2* er allerede isolert (Christin Vanberg, hovedoppgave UIO, 2001). Ved å sammenligne ekspresjonen av de to operonene vil man kunne få kunnskaper om blant annet genregulering og stressrespons hos *C. aurantiacus*.

Milepeler i oppgaven var:

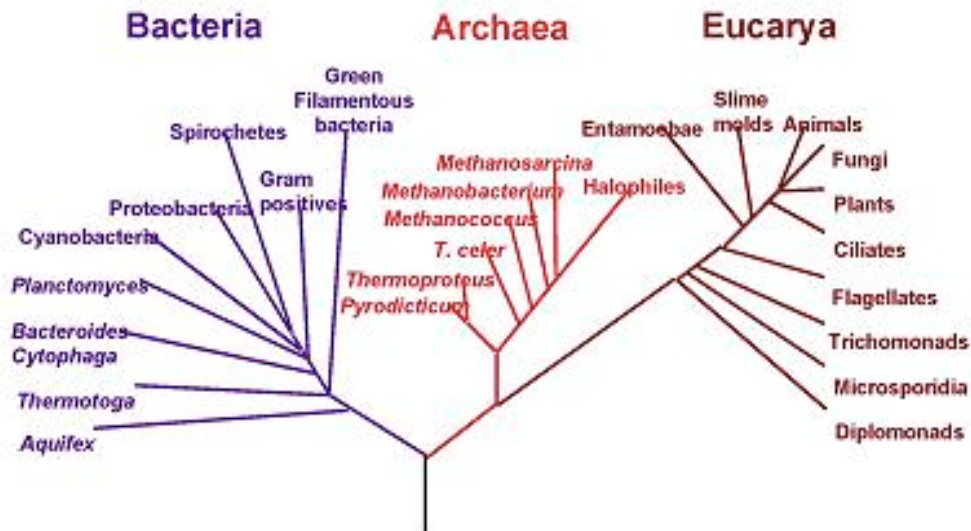
1. Å isolere hele operonet *groESL1* ved hjelp av den sekvensen som er tilgjengelig.
2. Å få uttrykket operonet *groESL1* i en ekspresjonsvektor.
3. Å undersøke det uttrykte proteinets aktivitet i nærvær av rensset malat dehydrogenase (MDH)
4. Å undersøke om GroESL fra *C. aurantiacus* er i stand til å gjøre *E. coli* mer termostabil.

FORKORTELSER

APS	ammoniumpersulfat
ATP	adenosintrifosfat
bp	basepar
BSA	bovin serum albumin
CIRCE	"controlling inverted repeat of chaperone expression"
dH ₂ O	destillert vann
DNA	deoksyribonukleinsyre
dNTP	deoksyribonukleosid 5`-trifosfatnukleotid
E. coli	Escherichia coli
EDTA	etylen-diamin-tetra-eddiksyre
HrcA	"heat regulation at CIRCE"
HSP	heat shock protein, varmesjokk-protein
IPTG	isopropyltio-β-D-galaktosid
kb	kilobasepar
kDa	kilodalton
L	liter
LB	Luria Bertani
MDH	malat dehydrogenase
min	minutter
ml	milliliter
mM	millimolar
μl	mikroliter
μg	mikrogram
μM	mikromolar
mRNA	messenger RNA
M _w	molekylvekt
Ng	nanogram
OD	optical density
O-mix	orange-mix
ON	overnatt
PCR	"polymerase chain reaction"
RNA	ribonukleinsyre
rpm	rotors pr. minute=omdreiningen per minutt
SDS	"sodium dodecyl sulphate"
SDS-PAGE	"sodium dodecyl sulphate polyacrylamide gel electrophoresis"
TAE-buffer	Tri-eddiksyre-EDTA-buffer
TE-buffer	Tris-EDTA-buffer
T _m	smeltetemperatur
TRNA	transport-RNA
uv	ultrafiolett
V	volt
v/v	vekt/volum

1 INNLEDNING

1.1 Livets tre



Figur 1.1. Fylogenetisk tre etter Carl Woese (1990). Kun 9 grupper i Bacteria er vist, i dag er 17 grupper kjent.

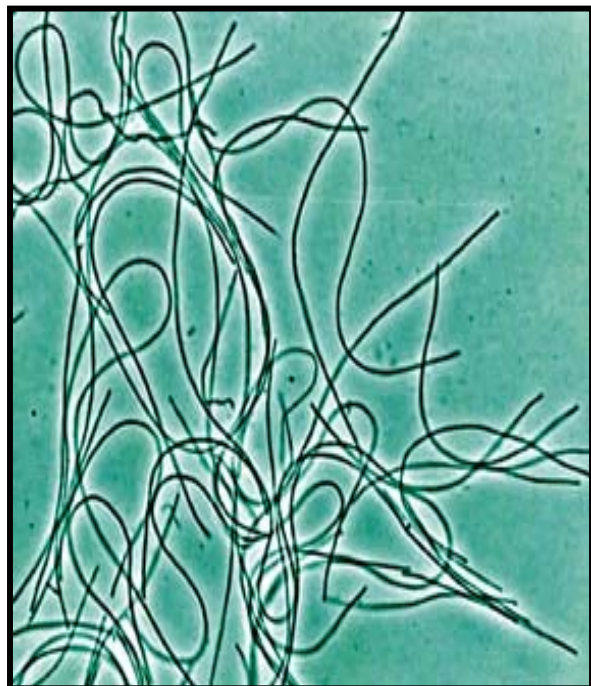
En måte å klassifisere organismene på, og som er mye brukt, er basert på analyse av 16S rRNA. Ved hjelp av denne metoden er et såkalt universelt fylogenetisk tre konstruert (Woese og Fox, 1977; Woese et al., 1990), der alt levende på jorden er delt inn i 3 store domener, Eukarya, Bacteria og Archaea (figur 1.1). Molekylet 16S rRNA er en del av ribosomene, og tilstede i alle levende organismer. Dette molekylet har en meget langsom evolusjonær utvikling, og forandringene i baserekkefølgen skjer sakte. En stor forandring i baserekkefølgen medfører stort sett celledød, mens langsomme, akkumulerte små forandringer tillater organismen å leve og formere seg. Ved å sammenligne 16S rRNA et hos ulike organismer kan man finne ut det evolusjonære slektskapet mellom dem. Jo likere rRNA organismene har, desto nærmere beslektet tror vi disse organismene er. Det fylogenetiske treet er basert på denne metoden. Avstanden på grenene som skiller organismene representerer den evolusjonære avstanden mellom dem.

Organismen i denne oppgave er medlem av riket Bacteria, der den tilhører gruppen av fototrofe bakterier. De fototrofe prokaryotene er en mangfoldig samling av organismer som har den felles egenskapen at de kan bruke lys som energikilde for vekst. Tilstedeværelsen av klorofyll, bakterieklorofyll, karotenoider, fykobiliner, bakteriens assosierte reaksjonssentere og lyshøstende proteinkomplekser skiller fototrofer fra ikke-fototrofer. Det er ingen universell akseptert avstammingslinje for alle dagens fototrofe prokaryoter, og fotosyntesens opphav og evolusjon er heller ikke kjent. De fototrofe prokaryotene er tilstede i ulike fylogenetiske grupper innenfor Bacteria og kan være mer beslektet med ikke-fototrofe slektinger enn med andre fotosyntetiske. Klassifiseringen av de fototrofe bakteriene baserer seg på fysiologiske forskjeller som for eksempel hvilke pigmenter og fotosystemer som er tilstede, om det produseres O_2 eller

ikke, hva som benyttes som elektronkilde for biosyntese osv. Av de sytten gruppene av Bacteria som er kjent, er det fem som inneholder fototrofe bakterier. Disse gruppene er purpurbakteriene med purpur svovelbakterier som slekten *Chromatium* og purpur ikke-svovelbakterier som slekten *Rhodobacter*, de grønne svovelbakteriene med slekten *Chlorobium*, cyanobakterier, Gram positive bakterier med *Heliobacterium* og de grønne glidende bakteriene med slekten *Chloroflexus*. Det er den grønne glidende bakterien *Chloroflexus aurantiacus* som er brukt i denne oppgaven. *Chloroflexus* er ulik de grønne svovelbakteriene og purpurbakteriene både i fylogeni og fenotyp. Visse likhetstrekk er allikevel å finne. *Chloroflexus* ligner purpurbakteriene ved at de har lignende fotokjemiske reaksjonssentere og ligner de grønne svovelbakteriene ved at de inneholder klorosomer og er fototrofe. De ligner den grønne svovelbakterien *Chloroherpeton* mest fordi de beveger seg med en glidende bevegelse. De grønne glidende bakteriene er trådformete organismer som har utviklet seg som en separat fylogenetisk gruppe, og finnes som en egen gren på treet, forholdsvis nær roten.

Fire grupper av grønne glidende bakterier er nå kjent, *Chloroflexus*, *Oscillochloris*, *Heliobacterium* og *Chloronema*. De siste to har ennå ikke blitt isolert i renkulturer. I tillegg omfatter denne fylogenetiske gruppen av bakterier den heterotrofe glidende bakterien *Herpetosiphon*. Familien Chloroflexaceae inneholder foreløpig tre arter, *Chloroflexus aggregans*, *Chloroflexus aurantiacus* og *Chloroflexus sp.*

1.1.1 *Chloroflexus aurantiacus*



Chloroflexus aurantiacus (figur 1.2) er den best undersøkte av de grønne glidende bakteriene. Det er en tynn, lang (0,5-1,0 μm i diameter, 30-300 μm lang) filamentdannende, glidende bakterie. Filamentene er uforgrenet og uten flageller. Bakterien ble første gang isolert fra varme, alkaliske kilder i Japan av Pierson og Castenholz i 1974.

I naturen finnes *C. aurantiacus* som lange tråder i såkalte mikrobielle matter. Voksetemperaturen varierer fra 40-70 °C, med en optimumstemperatur på rundt 55 °C. Dette plasserer *C. aurantiacus* i gruppen som en moderat termofil organisme. *C. aurantiacus* er funnet ved høyere temperaturer enn noen annen anoksygen fototrof. Den øvre temperaturgrensen for fotosyntese ser ut til å være rundt 70 °C (Meeks og Castenholz, 1971).

Figur 1.2. *Chloroflexus aurantiacus*

C. aurantiacus kan vokse under svært ulike forhold. Den kan vokse heterotroft ved aerob respirasjon, ftoheterotroft med lys som energikilde og organisk materiale som karbonkilde og fotoautotroft med lys som energikilde

og CO₂ som karbonkilde. Den vokser raskest som fotoheterotrof, anaerobt i lys. Den kan bruke en rekke organiske karbonkilder, som sukker, aminosyrer og organiske syrer til fotoheterotrofisk vekst. *Chloroflexus aurantiacus* kan også vokse kjemoheterotroft aerobt i mørket på organiske forbindelser (Madigan et al., 1974). Den får da tomatrød farge pga. manglende klorofyll og store mengder karotenoider. Endelig kan den også vokse fotoautotroft anaerobt i lys, med svovelforbindelser eller hydrogengass som elektrondonor. Disse egenskapene gjør denne bakterien meget fleksibel og tilpasningsdyktig til forandringer i miljøet. Som ekstern elektrondonor ved fotoautotrof vekst kan både molekylært hydrogen (Holo og Sirevåg, 1986) og sulfid, H₂S, (Madigan og Brock, 1975) brukes. Organiske komponenter brytes ned via TCA syklus (Sirevåg og Castenholz, 1979), men når acetat er eneste karbonkilde blir glykosylatsyklusen brukt (Løken og Sirevåg, 1982). Det er interessant å legge merke til at *C. aurantiacus* verken bruker Calvin-Benson syklusen eller den reduktive TCA syklusen til CO₂ fiksering men at den istedenfor bruker den unike 3-hydroxypropionat veien (Holo og Grace, 1987; Holo 1989; Stauss og Fuchs, 1993; Eisenreich et al., 1993). Denne veien er også kjent hos noen av Archaea, blant annet hos *Sulfolobus metallicus* (Burton, Williams, Norris, 1999.). I 3-hydroksypropionat syklusen reduseres 2 CO₂ molekyler til glyoksylat. Hvordan glyoksylat omdannes videre er enda ikke kjent (Ivanovsky og Krasilnikova, 1995).

Chloroflexus vokser bl.a. i tykke matter i nøytrale til alkaliske varme kilder der den er å finne i de nedre lagene med cyanobakterier som *Synechococcus* og *Mastigocladus*, i lagene over. I kilder med høyt svovelinhold, kan *Chloroflexus* finnes alene. *Chloroflexus* vokser vanligvis fotoheterotroft, antakeligvis på de organiske produktene til de autotrofe cyanobakteriene i de mikrobielle mattene (Bauld and Brock, 1974).

Det lyshøstende apparatet består av klorosomer festet til cellemembranen. Klorosomene hos *Chloroflexus* er noe mindre enn klorosomene hos grønne svovelbakterier. Klorosomene inneholder bakterieklorofyll Bchl c, mens det lyshøstende komplekset inneholder Bchl a, det samme som hos purpurbakteriene. Organismen har tre absorpsjonstopper i spekteret, ved 740 nm som skyldes Bchl a og ved 808 og 868 nm som begge skyldes Bchl c. Alle tre toppene er nær det infrarøde området i absorpsjonsspekteret. I tillegg til bakterieklorofyll inneholder *C. aurantiacus* karotenoidene β og γ karoten og OH-γ-karoten glukosid estere. Det fotokjemiske reaksjonssenteret er likt det vi finner hos purpurbakteriene, og inneholder pheophytin og quinon.

C. aurantiacus er en interessant bakterie å bruke i forbindelse med studier av fotosyntesesystemenes evolusjon. Fotosyntetiske eukaryoter er, sett i lys av den evolusjonære utviklingen, et relativt nytt fenomen, og forklares ofte ved hjelp av den såkalte endosymbiontteorien. Hovedtrekkene i denne teorien er at ikke-fotosyntetiserende eukaryoter tok opp i seg fotosyntetiske organismer. I denne sammenhengen er *Chloroflexus* av interesse når det gjelder å lete etter opphavet til type II fotosyntese reaksjonssenter. En organisme som *Chloroflexus*, som kan leve enten ved hjelp av respirasjon eller ved fotosyntese, egner seg til å spore fotosyntesens evolusjon.

Ifølge NASA er astrobiologi studier på ekstraterrestrialt liv. I 1996 ble dette fagområdet styrket. Fordi *Chloroflexus* er så mangfoldig og kan vokse under svært varierte forhold, har NASA valgt denne bakterien som en av modellorganismene for liv på andre planeter. Andre modellorganismer brukt i astrobiologien er cyanobakterien *Synechocystis* sp., den grønne svovelbakterien *Chlorobium tepidum*, proteobakterien *Rhodobacter capsulatus* og heliobakterien *Heliobacillus mobilis* (Blankenship, 2002, Xiong, 2000, Allen, 2000).

1.2 Termofile organismer

Temperatur er en viktig miljøfaktor, og organismer varierer stort i sitt forhold til å tilpasse seg denne faktoren. Da forskere på begynnelsen av 60-tallet begynte å studere varme kilder, spesielt i Yellowstone National Park, ventet ingen å finne hypertermofile bakterier. På grunn av den kunnskap en hadde på den tiden om virkning av varme på proteiner og nukleinsyrer, var det antatt at den øvre grensen for liv var omkring 73 °C. Det hadde vært kjent i over 100 år at enzymer ødelegges ved koking. Feltarbeidet viste at dette ikke var tilfelle, bakterier var å finne ved mye høyere temperaturer. Etter hvert som de varme kildene ble undersøkt, dukket det opp bakterier som kunne leve ved mye høyere temperaturer enn det som var antatt mulig. Disse oppdagelsene hadde stor betydning for videre arbeid innenfor biologien.

Alle organismer har en minimum, en maksimum og en optimumtemperatur. Kommer temperaturen over maksimumsgrensen, vil proteiner, nukleinsyrer og andre cellekomponenter degraderes. Lavere temperaturer enn minimumsverdiene kan være skadelige ved at membranene til cellen stivner og transport av næringsstoffer og dannelsen av protongradient på den måten blir umulig. Det er antatt at den nedre grense for liv er -12 °C, da det ved denne temperaturen dannes intracellulære iskrystaller (Russell, 1992). Trolig blir også kovalente bindinger for sterke ved lavere temperaturer, slik at reversible reaksjoner ikke lenger er mulig.

Det finnes få Eukaryoter ved høye temperaturer, den øvre grensen for disse er omkring 60–62 °C, med sopp og alger som representanter. Den øvre temperaturgrensen for planter og dyr er enda lavere, under 50 °C. Ved temperaturer over 60–62 °C finner man bare prokaryote organismer, både fra Archaea og Bacteria. Fotosyntetiske bakterier har en øvre temperaturgrense på 70–73 °C som er lavere enn den for de ikke-fotosyntetiserende. Når temperaturen blir over 100 °C, er det for de fleste organismer fra Archaea, de såkalte hypertermofile organismene, som er utbredt.

Termofile organismer finnes over alt på jorden, i varme kilder, vulkaner, ørkener, kraftverk, varmtvannsbeholdere osv. Prokaryotene finnes over hele området av temperaturer, men ingen av dem kan vokse ved alle temperaturer. Noen kan vokse i både mesofile og termofile omgivelser, og kalles fakultativ termofile. *Bacillus coagulans* er en slik bakterie med en temperaturløselighet på 30–60 °C.

Mange av de prokaryotene vi kjenner og som vokser under ekstreme forhold er medlemmer av Archaea. Mange ekstremofile organismer er trolig ennå ikke oppdaget, på grunn av vanskelig feltarbeid og problemer med å få dem til å vokse i laboratorier. Eksempler på ekstremofile organismer er *Pyrodicticum occultum* som har en optimumstemperatur på 105 °C, med minimumtemperatur på 82 °C og en maksimumstemperatur på 110 °C (Stetter, 1982) og *Pyrolobus fumarii*, med en maksimumstemperatur på 130 °C (Blöchl et al., 1997) som er å finne i områder rundt undersjøiske vulkaner. Den best undersøkte hypertermofile slekten er *Sulfolobus*. Denne archaean har en optimumstemperatur på rundt 80 °C og en optimum pH på rundt 2. Organismen lever i varme kilder og får energi ved å oksidere svovel. Mange av organismene som lever ved temperaturer nær eller over vannets kokepunkt er funnet i vulkanske områder, på land eller på havbunnen (Stetter, 1996), der de drar nytte av reduserte forbindelser som for eksempel fra vulkanene.

Hvordan er det mulig for disse organismene å overleve slike høye temperaturer? Noen av de termofile lever i temperaturer nær eller over vannets kokepunkt. Nukleinsyrer i løselighet i denne temperaturen vil raskt degraderes.

Hvordan holder termofilenes DNA seg stabilt? Hvorfor bryter ikke membranen sammen? Hvorfor denaturerer ikke proteinene? Dette er viktige og interessante spørsmål som mange forsøker å finne svar på.

1.2.1 Termostabilitet

En viktig begrensende faktor for liv ved høye temperaturer har å gjøre med membranens permeabilitet. Biologiske membraner blir lekk ved høye temperaturer og tillater fri passasje av ioner (Driessen et al., 1996). Dette må organismen unngå dersom ionegradienten over membranen skal kunne opprettholdes. For at en membran skal kunne fungere er det absolutt nødvendig at den er flytende. Fluiditeten til membranen er hel avhengig av fettsyrene i fosfolipidene. Mettede fettsyrer danner sterkere hydrofobe interaksjoner og har høyere smeltepunkt enn de umettede fettsyrene og gjør dermed membranen mer stabil og mindre flytende enn umettede fettsyrer. Det er vist at bakterier som vokser ved høye temperaturer har et høyere innhold av mettede fettsyrer enn bakterier som vokser ved lavere temperaturer (Oshima, 1978). Smeltepunktet øker også med lengden av fettsyrene. I *C. aurantiacus* er det påvist flere lange fettsyrer som kan være viktige for å holde membranen flytende (Knudsen et al., 1982; Shiea et al., 1991). Hypertermofile arter som tilhører Archaea har ikke fettsyrer men hydrokarboner i membranen. Hydrokarbonene kan ha forskjellige lengder og består av gjentatte enheter av phytan som er bundet til glyserolfosfat ved hjelp av en eterbinding.

DNA dobbelheliksen er også stabil ved økende temperaturer, opp til 107 °C, så lenge de to trådene ikke kan rotere fritt rundt hverandre. Dette er avhengig av de intracellulære forholdene (Marguet & Forterre, 1994). Virksomme DNA reparasjonsmekanismer er kritisk hos hypertermofile organismer. I forhold til DNA er RNA meget ustabil ved høye temperaturer på grunn av den reaktive OH-gruppen i 2' posisjon på ribosemolekylet. Det har blitt beregnet at et RNA molekyl på 2000 nukleotider blir delt i to etter 2 sekunder ved 110°C (Forterre, 1995). Områder på RNA som er utsatt for varmedestruksjon ser ut til å være beskyttet hos termofile tRNA og mRNA molekyler ved metylering av de reaktive OH-gruppene. Et høyt innhold av G og C i nukleinsyrene øker antall hydrogenbindinger og øker dermed stabiliteten. Det er imidlertid også påvist høyt G/C innhold i mesofile bakterier, og man kan derfor ikke si at det er direkte sammenheng mellom organismens G/C % og graden av termofili. Nukleinsyrer kan også stabiliseres av ytre faktorer, f.eks virker polyaminer stabiliserende på DNA fra *C. aurantiacus in vitro* (Norgaard, 1984).

Termostabiliteten til et protein sier noe om den evnen det har til å beholde sin tredimensjonale struktur ved høye temperaturer. Proteinet beholder strukturen bare innenfor et bestemt temperaturområde. Utenfor dette området blir proteinet denaturert og kan ikke lenger fungere. Noen proteiner fra termostabile organismer er fortsatt aktive *in vitro* etter inkubering ved 140 °C. For å oppnå en slik termostabilitet i proteinene ser det ut til at antall ikke-kovalente bindinger mellom aminosyrene som opprettholder den tredimensjonale strukturen til proteinet har økt gjennom evolusjonen.

Sammenligning av aminosyresekvenser og tredimensjonale strukturer til homologe proteiner fra mesofile og termofile bakterier har vist at det ikke finnes klare regler som kan forklare termostabiliteten til proteiner. Det ser ut til at de ulike enzymene har ulike løsninger på det å være termostabile. Det finnes imidlertid noen ting som peker i samme retning. Hydrogenbindinger er viktig for foldning av proteiner og det er vist at økt antall hydrogenbindinger øker

termostabiliteten til proteiner (Albert et al., 1987; Shortle, 1992; Rose og Wolfenden, 1993; Vogt et al., 1997; Wallon et al., 1997). Termofile proteiner har generelt mer kompakt sekundærstruktur enn homologe mesofile proteiner. Dette kan oppnås på flere forskjellige måter. Ved å minske fleksibiliteten til proteinkjedene eller forkorte løkkene, stabiliseres α -helikser. I termostabile proteiner er løkkene generelt kortere enn i mesofile homologer (Dagett og Levitt, 1993; Wallon et al., 1997). Fleksibiliteten avtar ved økt antall prolinenheter (Matthews et al., 1987; Herning et al., 1992). Prolinenheter i løkkeområder fører til krappere og stivere svinger. Det er vist et høyere antall prolin i løkkene hos termostabile proteiner enn hos mesofile (Russel og Taylor, 1995; Duffield et al., 1994). Det er også vist at termostabiliteten til et protein kan forbedres ved å sette inn prolinenheter i løkkene (Hardy et al., 1993) og at α -heliksene kan stabiliseres ved å minske fleksibiliteten. I termofile proteiner har α -heliksen generelt mindre fleksibilitet enn hos de mesofile. Fleksibiliteten kan gjøres mindre ved å øke antall alanin residier i heliksområder (Argos et al., 1979; Menendez-Arias og Argos, 1989; Blaber et al., 1995; Madeco-Ribeiro et al., 1997). Alanin fører til stivere og tettere pakket α -helikser med økt mulighet for hydrofobe kontakter. Sammenligning av primærstrukturen til homologe mesofile og termostabile proteiner har vist at i de termostabile proteinene er bestemte aminosyrer byttet ut med andre i forhold til de mesofile enzymene (Argos et al., 1979; Menendez-Arias og Argos, 1989). Aminosyrene som er mest foretrukne i de termostabile enzymene er aminosyrer som stabiliserer heliksstrukturen, øker intern hydrofobisitet, stabiliserer β -plater og gir bedre pakking.

Den første delen av varmeinaktivering er denaturering av molekylet. Under normale forhold blir den native, katalytiske aktive strukturen til proteinet holdt i balanse ved hjelp av forskjellige ikke-kovalente krefter som for eksempel hydrogenbindinger, hydrofobe, ioniske og van der Waalske interaksjoner. Når temperaturen stiger, kan økt fri energi forårsake forandring i strukturen fra nativ katalytisk aktiv form til en tilfeldig foldet inaktiv form. I denne prosessen blir lavenergibindinger som stabiliserer den tredimensjonale strukturen ødelagt for eksempel på grunn av økt vibrasjonsenergi eller på grunn av kollisjon med vannmolekyler. Kollisjon med vannmolekyler øker når temperaturen øker (Ward og Moo-Young, 1988). Ved denaturering blir hydrofobe områder til proteinet eksponert i løsningen. For å minske en slik eksponering vil proteiner binde seg sammen med hverandre i hydrofobe regioner. Slike interaksjoner resulterer i proteinaggregering og forekommer særlig i konsentrerte proteinløsninger.

Hva er det som gjør et protein mer stabilt enn et annet? Hittil er det ikke funnet klare entydige svar på dette spørsmålet, men det er kjent at både den interne strukturen til proteiner og faktorer i omgivelsene bidrar til termostabilitet. Proteiner fra termofile organismer egner seg svært godt til å undersøke proteinestabilitet. Innsikt i proteinenes termostabilitet er av generell og bioteknologisk interesse. I industrielle prosesser kan høye temperaturer gi fordeler som økt løselighet av substrat og minsket risiko for forurensninger av bakterier. Bruken av termofile enzymer som degraderer polymerer, som for eksempel amylaser, proteaser og cellulaser, kan spille en betydelig rolle i mat industrien, kjemisk industri, farmasøytisk industri og avfallsindustrien (Schiraldi og De Rosa, 2002). I termofile bakterier er det funnet unike termostabile enzymer. Rensete proteiner fra termofiler viser ofte en høyere stabilitet ved normale temperaturer, en egenskap som har tiltrukket seg oppmerksomhet fra kommersielt hold. Forskning på ekstreme miljøer har et stort potensiale når det gjelder bioteknologisk forskning. Eksemplene er mange. *Bacillus stearothermophilus*, med en maximumstemperatur på 75 °C, blir brukt i industrien til å produsere enzymer til vaskepulver. *Thermus aquaticus* varmetolerante polymerase revolusjonerte PCR metoden.

Foldningskatalysatorer, kjent som chaperoner, syntetiseres ofte i økt mengde i cellene under varmesjokk. De sørger for korrekt foldning av polypeptider og forhindrer aggregering av ufoldede proteiner (Trent, 1996). Det er vist at chaperonene stabiliserer proteinene (Hartman et al., 1993; Ranson et al., 1995). Det er varmesjokkresponsen og chaperoniner som er temaet for denne hovedoppgaven, og det er dette de neste avsnittene skal dreie seg om.

1.2.2 Varmesjokkrespons

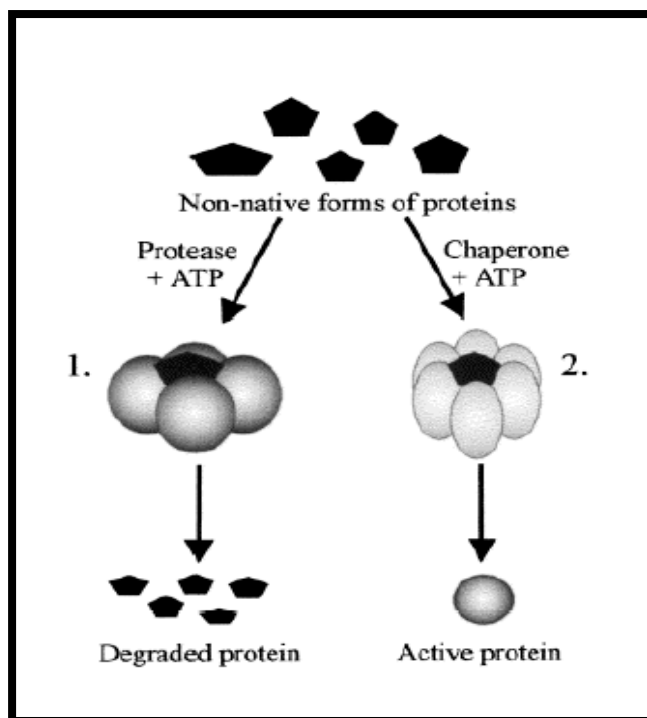
Stress for en bakterie er forandringer som fører til redusert veksthastighet, forandringer i miljøet som resulterer i skader på cellen og situasjoner som stimulerer uttrykk av gener som er forbundet med spesifikke forandringer i miljøet. Stress fører til at cellen må gjennomgå tilpasninger av ulike slag for å kunne overleve. Eksempler på stressfaktorer er varme, kulde, pH, stråling og oksygentrykk. Det finnes en rekke overlevelsesmekanismer mot ulike grader og konsentrasjoner av disse faktorene. Ekstremofile bakterier for eksempel lever under forhold ikke mulig for andre å vokse i. Uansett miljø er organismen tilpasset, og skulle det oppstå en plutselig forandring i miljøet, må den ha en respons på forandringen for å kunne overleve, en såkalt stressrespons.

Når en celle er under stress, finner det sted en rekke ulike forandringer på transkripsjonsnivå, mRNA syntesen, mRNA stabiliteten og translasjonseffektiviteten blir påvirket. For å beskytte seg mot forandringer forårsaket av stress, er det nødvendig for en celle å tilpasse proteinsammensetningen, for eksempel ved å syntetisere såkalte varmesjokkproteiner. Denne meget velregulerte forandringen i transkripsjon kalles varmesjokkrespons. Til tross for navnet er dette en respons på ulike typer stress, ikke bare økning i temperatur, derfor kalles det også stressrespons.

"Heat Shock" Proteiner (HSPs), også kalt stressproteiner, er en gruppe proteiner som har til felles at transkripsjonen av dem oppreguleres (induseres) dersom cellen utsettes for stress. Denne responsen kalles varmesjokkrespons og finner sted i alle organismer hittil undersøkt. Den generelle funksjonen til disse proteinene er å forhindre aggregering og feilfoldning av andre proteiner. Når en celle utsettes for varmestress har proteinene i cellen vanskeligere for å oppnå riktig konformasjon og struktur på grunn av høyere kinetisk energi i polypeptidkjeden. Dette fører ofte til feilfoldning av proteinene og til at proteiner klumper seg sammen til uløselige aggregater, aggregering. Systemer som aktivt opprettholder og kontrollerer proteinstruktur er derfor en forutsetning for at cellen skal kunne overleve. Systemene involverer chaperoner og proteaser (Wickner et al., 1999), som er involvert i henholdsvis refolding og nedbrytning (Herman og D'Ari, 1998) (figur 1.3).

HSPs er også tilstede i cellene under normale forhold, der de sørger for at proteinene får riktig form, at de er tilstede på riktig sted i cellen, til riktig tid og i riktig konsentrasjon. Misdannede proteiner hjelpes tilbake til sin riktige form, proteiner transporteres til riktig sted, "gamle" og ødelagte proteiner transporteres til destruksjon. Det er mulig stressproteinene spiller en rolle i immunsystemet til pattedyr der de antageligvis hjelper til med å presentere peptider i cellemembranen på infiserte celler.

I en stresset celle kan HSPs utgjøre opp til 20 % av cellens løselige proteiner, mens de i celler under normale forhold, utgjør omkring 2 %. Når *Pyrodictum occultum* dyrkes ved 108 °C, utgjør omtrent 80 % av løselige proteiner fra cellefritt ekstrakt chaperonin-lignende proteiner (Phipps et al., 1991; Phipps et al., 1993). Når varmesjokk-responsen er indusert, kan kulturer av denne organismen overleve 1 time autoklavering ved 121 °C (Hoffman, 1993).



Figur 1.3. Skjematisk fremstilling av denaturerte proteiner, som enten degraderes av proteaser eller tilbakefoldes av chaperoner (Tjakko, A.; Wouters, J. A. 199)).

Uttrykking av varmesjokk-gener hos bakterier kontrolleres på transkripsjonsnivået av både positive og negative mekanismer. Positiv kontroll innebærer bruk av alternative sigmafaktorer som leder RNA polymerasen til varmesjokk promotoren. Negativ kontroll innebærer repressoravhengige mekanismer hvor transkripsjon av varmessjokk-gener blir satt igang fra en klassisk husholdningspromotor, og hvor cis-virkende DNA elementer inngår sammen med et tilhørende repressorprotein for å begrense transkripsjon under fysiologiske betingelser.

E. coli er den bakterien der varmesjokkrespons er best undersøkt. Den har mer enn 20 ulike gener for HSPs. Regulering hos *E. coli* er positiv, og basert på alternative σ faktorer. Responsen her er en økt ekspresjon, økt stabilitet og økt aktivitet av tre alternative σ -faktorer av RNA polymerase, kalt sigma 32, sigma 24 og sigma E. Disse binder seg med høy affinitet til oppstrømsregionene for genene til HSPs og andre stressrelaterte proteiner og blir derfor transkribert i en større grad når gener for de alternative σ -faktorene er tilstede og aktive, enn når disse ikke er tilstede. Denne måten for regulering er også kjent fra en rekke andre bakterier. Sigma 32 er kodet av *rpoH* og kontrollerer transkripsjon av alle varmesjokkproteiner i cytoplasma, inkludert chaperonene GroE og DnaK samt ATP-avhengige proteaser som Clp og Lon (Gross, 1996; Yura et al., 1993). Under varmesjokk øker nivået av σ 32 raskt både på grunn av økt translasjon av *rpoH* og mRNA, og på grunn av midlertidig økt stabilisering av σ 32. Dette fører i sin tur til økt transkripsjon av varmesjokkgenene. Responsen er "feedback" kontrollert av produktene til genene *dnaK*, *dnaJ* og *grpE*. Proteinene DnaK, DnaJ og GrpE samarbeider ved å presentere σ 32 til FtsH protease som bryter ned sigmafaktoren (Herman et al., 1995; Tomoyasu et al., 1995). Mens DnaK og DnaJ samarbeider for å hemme aktiviteten til sigma32 ved å forhindre at den binder til RNA polymerasen, reverserer delvis GrpE denne hemningen (Gamer et al., 1996). Under varmesjokk vil DnaK, DnaJ og GrpE bli opptatt med denaturerte proteiner, sigma32 vil dermed få økt stabilitet og aktivitet (Grossm, 1996). En oversikt over σ -faktorene hos *E. coli* er vist i tabellen nedenfor. Promotersekvensene varierer fra organisme

til organisme, men det er likheter mellom organismer i samme familie. *E. coli* har 7 hovedtyper promotersekvenser. Sigma 70 er den som vanligvis benyttes ved transkripsjon av de fleste genene (Kundu et al., 1997).

Sigmafaktor	type	-35	mellomsekvens -10	
$\sigma 70$	generell	TTGACA	16-18bp	TATAAT
$\sigma 54$	nitrogen	CTGGNA	6bp	TTGCA
$\sigma 38$ (σs)	stasjonær	ukjent	ukjent	ukjent
$\sigma 32$	heat shock	CCCTTGAA	13-15bp	CCCGATNT
$\sigma 28$ (σF)	flagell	CTAAA	15bp	GCCGATAA
$\sigma 24$ (σE)	heat shock	ukjent	ukjent	ukjent
$\sigma 19$ (σf_{ec})	jerntransport	ukjent	ukjent	ukjent

σ -faktorene og deres konsensussekvenser i *E. coli*. (Lewin, 2000; EcoCyc, 2004; Ussery, 1999)

Promotersekvensen for *groESL1* hos *C. aurantiacus* ligner på de vanlige husholdningssekvensene og sekvensene for sigmafaktorer som benyttes for varmesjokkgenene.

	-35 område	-10 område
Konsensus $\sigma 32$	CCCTTGAA	CCCGATNT
<i>C. aurantiacus</i> :	GGGTAGT GCGTGAC ACTACCGCCA CCCTGA ATAATGAGGA	
Konsensus $\sigma 70$	TTGACA	TATAAT

Studier av varmesjokkresponsen hos ulike bakterier har vist at positiv regulering slik som hos *E. coli* ikke er en universell reguleringsmekanisme.

Negativ regulering er basert på et spesifikt samspill mellom et repressorprotein og et DNA element som er lokalisert i promotorområdet til varmesjokk-genene. Et av de mer utbredte negative kontrollsystemene består av *HrcA* (Heat Regulation at CIRCE) og DNA elementet CIRCE (Controlling Inverted Repeat of Chaperon Expression). Kontroll av CIRCE-*HrcA* systemet ser foreløpig ut til å ligne på "feedback" kontrollen av sigma 32 hos *E. coli*, hvor DnaK kontrollerer stabiliteten og aktiviteten til sigma 32.

CIRCE elementet er en DNA sekvens som er sterkt konservert hos en rekke fylogenetisk forskjellige organismer (Segal og Ron, 1996). Sekvensen er å finne i det oppstrømsregulerende området for en rekke HSPs. CIRCE elementet består, med små variasjoner i noen organismer, av følgende sterkt konserverte sekvens:

TTAGCACTC-N9-GAGTGCTAA

Det er lagt frem ulike hypoteser om funksjonen til CIRCE elementet. En hypotese er at CIRCE virker stabiliserende på mRNA fra det aktuelle genet, en annen er at elementet er plassert der et repressor protein bindes. Dette proteinet kalles *HrcA* som står for Heat Regulating at CIRCE. *In vitro* studier har vist at transkripsjonen av gener som har CIRCE elementet reduseres i nærvær av *HrcA*, ved fysiologiske forhold, mens *HrcA* ikke har noen virkning på gener uten CIRCE elementet. Nok en annen hypotese går ut på at GroESL har en viktig funksjon med å få *HrcA* i riktig konformasjon. I *Streptomyces* er *HrcA* funnet i et varmesjokk-operon, og blir derfor oppregulert ved varmesjokk. Fordi

GroESL komplekset er opptatt med å refolde proteiner som er feilfoldet på grunn av varmesjokket, vil ikke HrcA foldes riktig og bli aktivt før antall feilfoldete proteiner har sunket. HrcA foldes riktig av GroESL og bindes til CIRCE elementet, og på denne måten nedregulerer HSPs og operon, samtidig som HrcA selv nedreguleres. (Yura, et al., 1993).

Det er vist at noen bakterier inneholder flere varianter av gener for HSPs, blant annet av *groESL* operonet. Et spørsmål i forbindelse med at organismen har flere kopier av slike gener, er om genene har ulik regulering og/eller funksjon. *Bradyrhizobium japonicum* har 5 operoner for *groESL* og studier av disse har vist at kontrollen av dem involverer fire ulike elementer på transkripsjonsnivå (Fischer et al., 1993; Babst et al., 1996). Under normale vekstbetingelser er operonet *groESL2* uttrykt konstitutivt. Operonene *groESL1*, *groESL4* og *groESL5* er alle varmeindusert, men *groESL1* er under kontroll av sigma 32, *groESL4* og *groESL5* under negativ regulering av CIRCE elementet. Operonet *groESL3* er ko-regulert med nitrogenfikseringsgener via sigma 54 og transkripsjonsaktivator proteinet NifA (Fischer et al., 1993).

I *C. aurantiacus* er CIRCE elementet funnet i *groESL2*, men ikke i *groESL1*. Det er også funnet en HrcA homolog i genomet. Oppstrøms for HrcA genet, er det også funnet et CIRCE element. Dette kan tyde på at HrcA nedregulerer seg selv under ikke-varmesjokk-forhold. At de to genutgavene har ulike mekanismer for regulering kan bety at de transkriberes og aktiveres under ulike forhold av stress. Tilstedeværelse av flere varianter av *groESL* gener kan ha betydning ved at de gir bakterien fleksibilitet ved varierende miljøbetingelser og fysiologiske krav. Flere utgaver av de samme genene kan ha flere fordeler slik som økt ekspresjon av det gjeldende proteinet avhengig av forholdene bakterien befinner seg i og dessuten en viss forsikring mot mutasjoner eller andre ødeleggende faktorer. Et av målene med å studere de to genutgavene av *groESL* hos *C. aurantiacus* er for å få kunnskap om disse mekanismene.

1.2.3 Molekylære chaperoner

Chaperonene er en gruppe av varmesjokkproteinene. De er blant de mest konserverte proteinene i naturen og funnet i alle organismer som er undersøkt til nå, bortsett fra to arter av *Mycoplasma*.

Prosessen et polypeptid gjennomgår for å oppnå riktig tredimensjonale struktur og biologisk aktiv tilstand, kalles for proteinfolding. Anfinsen og medarbeidere utførte det klassiske forsøket som viste at all informasjon som var nødvendig for folding av polypeptidkjeder til riktig tredimensjonale struktur lå i rekkefølgen av aminosyrer (Epstein et al., 1963). Små polypeptider, som består av mindre enn 100 aminosyrer, kan folde spontant uten foldingsintermediater (Jackson, 1998), mens større polypeptider, mer enn 100 aminosyrer, vil folde seg gjennom forholdsvis stabile intermediater. Chaperoner kan binde og folde ny-syntetiserte polypeptider (Hendrick og Hartl, 1993). Under translasjon på ribosomene kan polypeptider med hydrofobe områder reagere med hverandre. En del polypeptider vil begynne å folde under translasjon, men for større proteiner med flere domener vil dette ikke være mulig. Nå en celle utsettes for stress, blir aggregering et enda større problem fordi allerede foldede proteiner begynner å denaturere og kan reagere både med hverandre og med nylagde proteiner. For å kunne takle disse problemene produserer alle celler de universelle, konserverte proteinene som vi kaller molekylære chaperoner. Hovedfunksjonen til disse proteinene er å minimalisere proteinaggregering og sørge for riktig foldning. Chaperoner binder og refolder feilfoldete og misdannede

proteiner og forhindre protein aggregering ved å binde til utsatte hydrofobe deler til ikke native polypeptider (Hartl, 1996; Bukau og Horwich, 1998). Nylagde proteiner kan dermed holdes i en gunstig tilstand for riktig foldning. Riktig foldning kan også sikres ved kontrollert binding og frigjøring av polypeptidet fra chaperonet. Hovedsakelig vil chaperoner sørge for trygge betingelser der polypeptidet kan folde under instruks fra aminosyresekvensen. Energien som brukes i foldelsesprosessen kommer fra hydrolyse av ATP, som fører til en konformasjonsforandring i chaperonet.

De molekylære chaperonene utgjør store proteinfamilier hvor mange er de såkalte varmesjokk-proteinene, HSP, som syntetiseres i store mengder når temperaturen øker (Nover, 1991). Chaperonene deles opp i flere klasser basert på molekylvekten, kDa, til proteinene med Hsp 100, 90, 70, 60, 40 og 27/28. Chaperonene Hsp60 og Hsp70 er de klassene som hittil er mest undersøkt og best forstått.

Chaperoner kan i tillegg ha andre funksjoner som å sørge for korrekt transport og nedbrytning av proteiner (Gething og Sambrook, 1992; Morimoto et al., 1994; Hartl, 1996). Chaperonene blir selv ikke forandret i foldingsprosessen og de har heller ikke med seg noen informasjon om korrekt foldingsstruktur. Alle molekylære chaperoner gjenkjenner hydrofobe deler av ufoldete polypeptider, og "diskriminerer" derfor mot foldete proteiner.

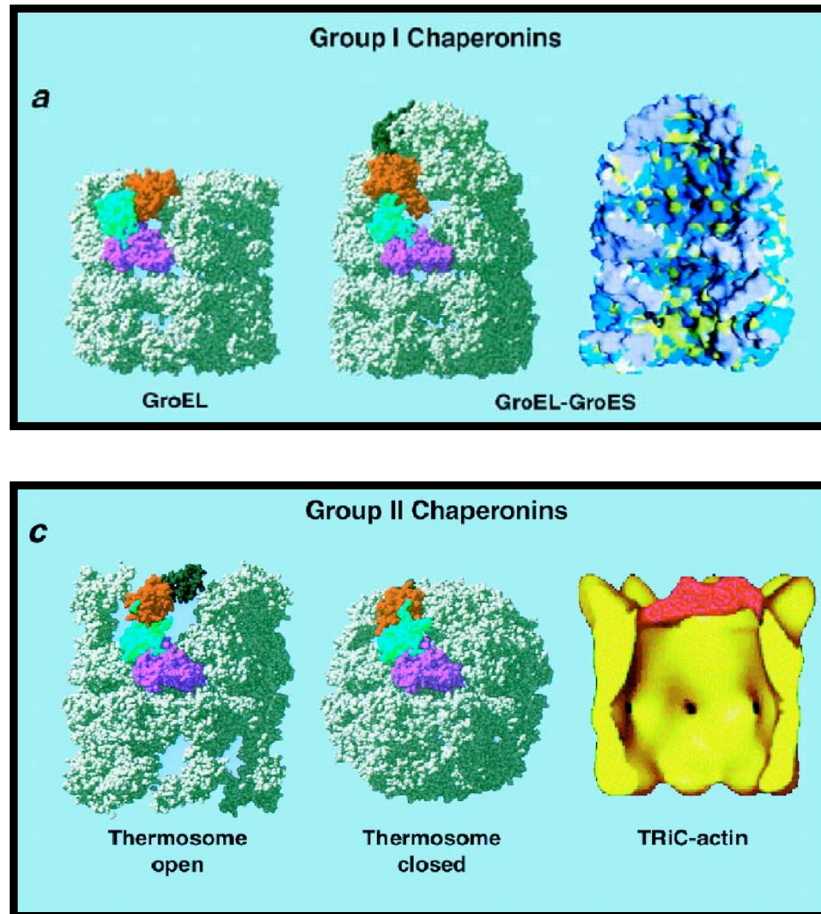
Mekanismen for prosessen via hjelpe-proteiner er bedre forstått hos bakterier der chaperoninene deltar i selve foldingsprosessen.

1.2.4 Chaperoniner

Det chaperonesystemet som er best undersøkt er chaperoninet Hsp60 og ko-chaperoninet Hsp10. Disse kalles også for chaperoniner og deres oppgave er å forhindre aggregering av proteiner og sørge for riktig folding. Chaperoninene kan folde og refolde strukturelt forskjellig proteiner (Horwich et al., 1993; Ewalt et al., 1997), og er nødvendige for at cellen skal overleve ved alle temperaturer (Fayet et al., 1989; Cheng et al., 1989). De uttrykkes konstituttivt, men mengden av dem øker sterkt ved stress slik som ved en plutselig økning i temperatur. Chaperoninene tar også del i nedbrytning av mange proteiner (Horwich et al., 1999). I en *E. coli* stamme med underskudd av GroEL er omtrent 30% av proteinene misdannet. Renset GroEL reagerer med mer enn 50% av de løselige *E. coli* proteinene hvis de er feilfoldet.

Basert på sammenligning av aminosyresekvenser er chaperoninene delt inn i to grupper, gruppe I og gruppe II (figur 1.4). Gruppe I chaperoniner finnes i alle eubakterier og i mitokondrier og kloroplaster hos eukaryoter. Chaperoninene i denne gruppen har en stor subenhet på omtrent 60kD (Hsp60, GroEL) og en liten subenhet, ko-chaperonin, på omtrent 10kD (Hsp10, GroES). Genene for disse to enhetene er vanligvis organisert i et operon, dvs. at de har et felles regulatorisk område oppstrøms for begge genene og blir derfor transkribert samtidig. Proteinene GroEL er multimer og består av to ringformete heptamerer som bindes sammen bunn mot bunn. GroES danner en liten heptamer, en lokk-lignende struktur, som fester seg på toppen av en av GroEL ringene (se figur s.30). Gruppe II chaperoniner har to store subenheter. I Archaea blir disse proteinene sterkt induisert av varme, og hos noen arter utgjør de hovedmengden av proteiner i cellen etter varmesjokk, de kalles her thermosomer, TRiC (Tailless Complex Polypeptide 1 (TCP1) Ring Complex (TRiC) og CCT (chaperonin containing TCP1). CCT finnes i cytosol hos eukaryote celler og hos

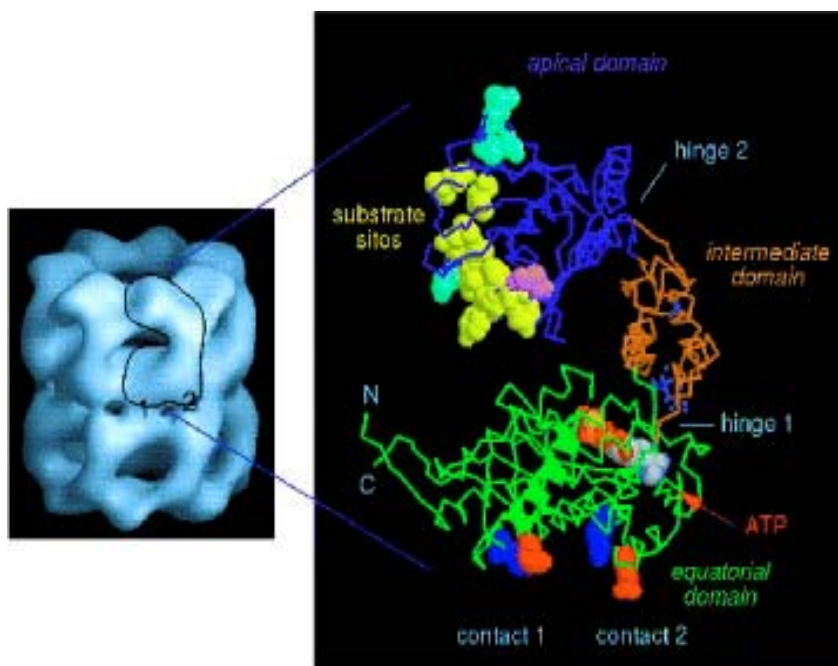
archaea. Den største forskjellen mellom type I og type II chaperoniner er formen på multimeren. Mens den store subenheten i type I chaperoniner danner to heptameriske ringer, danner type II chaperoninene to oktameriske ringer. I tillegg er type II ringstrukturen lukket på toppen, mens type I har en åpen ring på toppen der GroES bindes.



Figur 1.4. Bilde av de to gruppene chaperoniner.

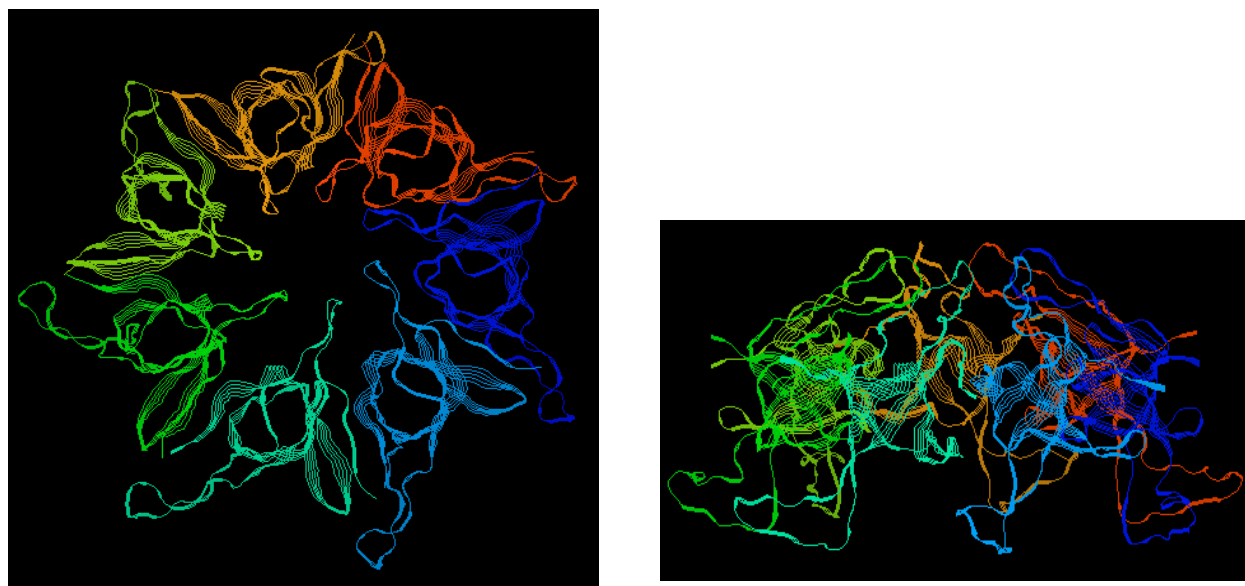
Både chaperoninet og co-chaperoninet er 7-foldete symmetriske oligomerer. Den store enheten, GroEL, er satt sammen av 14 identiske 60kDa subenheter ordnet i 2 stabler som 7-delte ringer som ligger oppå hverandre (figur 1.5). I hver GroEL ring danner de 7 subenhetene 3 domener oppover i ringen, apikalt, intermedialt og ekvatorial. Det ekvatoriale domenet er størst og består hovedsakelig av α -helikser og sørger for kontakt mellom de to ringene og mellom de enkelte subenhetene. Dette domenet er lokalisert i bunnen av ringstrukturen og er derfor det domenet som har mest intersubenhetskontakt. Det er også på dette domenet mesteparten av det ATP bindende setet er plassert. ATP bindes til ekvatorialdomenets ATP bindingseter og de C- og N-terminale endene. De C-terminale endene stikker innover fra hvert ekvatoriale domene og blokkerer midtgangen mellom hulrommene. Det ser dermed ikke ut til at polypeptider kan bevege seg fra et hulrom til et annet (Weissman et al., 1996; Rye et al., 1997). Det apikale domenet sitter på toppen av ringen, og danner åpningen på kanalen ned til det sentrale hulrommet. Dette området har høy fleksibilitet. Dette domenet inneholder hydrofobe aminosyreenheter som er involvert i bindingen av hydrofobe områder til ikke-native polypeptider og til GroES. Det intermediale domenet er lite og inneholder resten av det ATP bindende setet. Det

intermediale domenet har korte antiparallelle kjeder som virker som hengsler og er setet for rotasjon under konformasjonsforandringer i subenheten.



Figur 1.5. Struktur av GroEL ringkompleks og subenhet. Til venstre: krystallstrukturen til GroEL komplekset, 2,5 nm oppløsning (Braig et al., 1994) viser den dobbelt-ringede strukturen og plasseringen av en individuell subenhet innenfor oligomeren. Til høyre: modell av en enkel subenhet. Figuren er basert på Roseman et al., 1998, hentet fra Ranson et al., 1998, *Biochemical Journal* (1998) Volume 333.

Den lille subenheten i chaperonin komplekset kalles GroES og er på omtrent 10kD. Det danner en homoheptamerisk løkk-lignende struktur som intragerer med det apikale domenet til GroEL (figur 1.6). GroES er satt sammen av syv identiske 10kDa subenheter satt sammen i en ring. Hver subenhet har en beta tønnestruktur og to store løkkeområder som stikker ut fra hovedkjernen. En av løkkene stikker ut og oppover og danner et slags løst lokk over GroES ringen (Hunt et al., 1996). Det andre løkkeområdet er rikt på hydrofobe enheter og stikker ut og nedover for å danne kontakt med hydrofobe enheter i det apikale domenet til GroEL (Roseman et al., 1996; Landry et al., 1993).

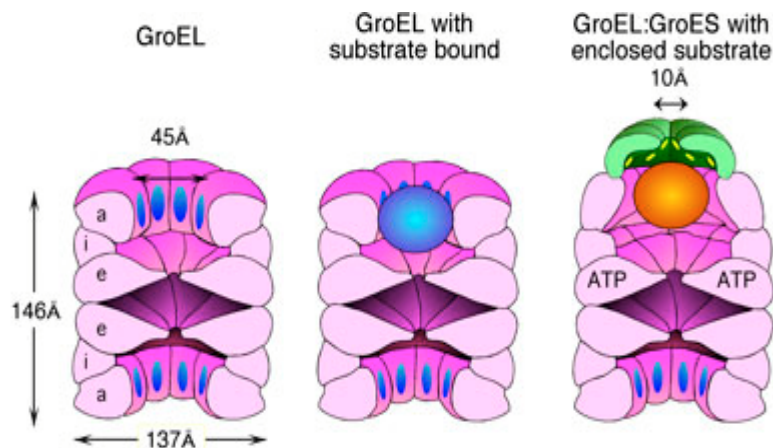


Figur 1.6. Struktur av GroEL, til venstre: komplekset sett ovenifra, til høyre: komplekset sett fra siden. Strukturen er rik i beta strukturen. Disse danner en antiparallell beta tønne-struktur, der innsiden er sterkt hydrofob. Toppen av strukturen dannes av hårnålstrukturer. Løkkene som strekker seg ned og ut fra lokket er svært mobile, men stabiliseres ved binding til GroEL komplekset.

Et ufoldet protein binder til seg den store hydrofobe overflaten inne i hulrommet i chaperonet (figur 1.7). Binding av ATP setter igang rotasjon av en Hsp60 subenhet slik at den hydrofobe overflaten vender seg mot nabo Hsp60 subenheten. Denne forandringen i konformasjon er koordinert med Hsp10, som lukker hulrommet med det ufoldete proteinet inni. Binding av ATP til det ekvatoriale domenet starter forandringen i konformasjonen i det apikale og intermediale domenet som fører til at det foldete proteinet løsner. Samtidig skjer det en konformasjonsforandring i GroEL som snur den hydrofobe overflaten bort fra hulrommet. Hulrommet blir hydrofilt. Dette frigjør det hydrofobe proteinet og slipper det fritt ut i hulrommet der det kan foldes til sin native struktur i forhold til informasjonen fra den primære aminosyrestrukturen. Frigjøring av det ferdigfoldete proteinet skjer når ATP og GroES binder seg til GroEL ringen på den motsatte siden, noe som utløser en ugunstig ring-ring interaksjon og fører til disosiasjon av den første GroES og frigjøring av proteinet.

Den doble ringen i GroEL utfører proteinfoldingen i det sentrale hulrommet, ved hjelp av ATP og bundet GroES, men det er enda uklart hvordan GroEL utfører syklusen fra et aktivt foldelseskompleks til den neste. De to ringene i chaperoninkomplekset har minst to distinkte domener kalt *cis* og *trans* (Martin et al., 1991; Badcoe et al., 1991). GroEL ringen som har bundet GroES er i *cis* konfigurasjon, den som ikke har GroEL bundet er i *trans* konfigurasjon. Hydrolyse av ATP inne i *cis*-ringen må skje før enten det ikke-native proteinet eller GroES kan bindes til *trans*-ringen. Dette er assosiert med reorienteringen av *trans*-ringens apikale domene. Dannelsen av en ny *cis*-orientert ring avhenger av den åpne *trans*-ringen med polypeptidbinding først, som stimuleres av ATP-avhengig dissosiasjon fra *cis*-komplekset, fulgt av GroES binding. Dette tyder på at GroEL alternerer ringene i komplekset i nærvær av feilfoldete proteiner, som foldings-aktive *cis*-komplekser. En alternering krever 7 ATP. Hydrolyse av ATP vil svekke bindingen av GroES, og *cis* komplekset vil dissosiere, men ikke før ATP har bundet til *trans* ringen (Kad et al., 1998). GroES har høy

affinitet for GroEL-ATP komplekset og vil binde på nytt til trans-siden. Binding av ATP og GroES forårsaker en strukturell forandring til et stabilt cis-kompleks med et utvidet hulrom. Konkurransen mellom cis og trans konformasjonene, med ATP som stabiliserer og dominerer over ADP, bestemmer dermed over konformasjonen (Sigler et al., 1998).



Figur 1.7. Det ufoldete proteinet fester seg til det hydrofobe området (blått) på det apikale domenet i hulrommet. GroES binder seg til GroEL cis-ringene, samtidig som 7 ATP bides til det ekvatoriale domenet. GroES bindingen inducerer konformasjonsforandringer i GroEL via intermediatdomenet. Dette resulterer i et skifte oppover og utover i tillegg til en vridning av det apikale domenet (med klokken). Denne forandringen fører til at de hydrofobe områdene inne i hulrommet flytter seg slik at hulrommet blir hydrofilt. Volumet til hulrommet er $\sim 175\,000\text{\AA}^3$ og kan romme polypeptider på rundt 60kDa.

Etter konformasjonsforandringen blir hulrommet i GroEL hydrofilt og delvis lukket fra omgivelsene. Inne i hulrommet vil polypeptidet bli frigjort og får anledningen til å foldes. Etter en periode på 6-8 sekunder vil ATP hydrolysere i cis-komplekset (Burston et al., 1995; Weissman et al., 1996; Rye et al., 1997). GroES dissosierer fra komplekset ca. 5-10 sekunder etter dette ved at ATP bindes til trans-ringene. Polypeptidet må foldes riktig på denne tiden, omtrent 15 sekunder. Polypeptider som er for store, eller av andre grunner ikke rekker dette, må gå gjennom syklusen en gang til (Todd et al., 1994; Burston et al., 1996; Weissman et al., 1994; Smith og Fisher 1995).

Det er en øvre grense for størrelsen på polypeptidene som kan foldes av chaperoninene GroEL. Utvendig er proteinet 137 Å i diameter og 177 Å bredt (Landry et al., 1998). Den maksimale størrelsen på hulrommet, etter binding av GroES og konformasjonsforandring i cis hulrommet, er diameteren 8nm, 80Å. Polypeptider større enn dette må foldes via andre molekylære chaperoner.

1.3 Isolering av gener ved bruk av PCR

PCR ble beskrevet første gang på en konferanse oktober 1985 av Kery Mullis som også tok patent på oppfinnelsen (Mullis, 1990). PCR står for Polymerase Chain Reaction og er en reaksjon som baserer seg på forlengelse av korte oligonukleotider, såkalte primere, for å oppformere spesifikke nukleinsyrer *in vitro*.

Selve PCR reaksjon deles inn i 3 trinn: denaturering av templatet til enkelttrådig DNA, annealing av primere til hver tråd og elongering av nye tråder fra primerene. Disse 3 trinnene styres av tid og temperatur og utgjør en syklus som gjentas 25-30 ganger. Syklusene fører til eksponentiell økning av et spesifikt DNA fragment der endene til fragmentene er definert av de 5'-endene til primerene. Produktene som syntetiseres i en syklus vil virke som templat i neste syklus og dette vil føre til at antall kopier av den ønskede DNA sekvensen øker eksponentielt. Etter n antall sykluser vil reaksjonen teoretisk sett inneholde 2^n DNA-molekyler som alle er kopier av det originale templatet.

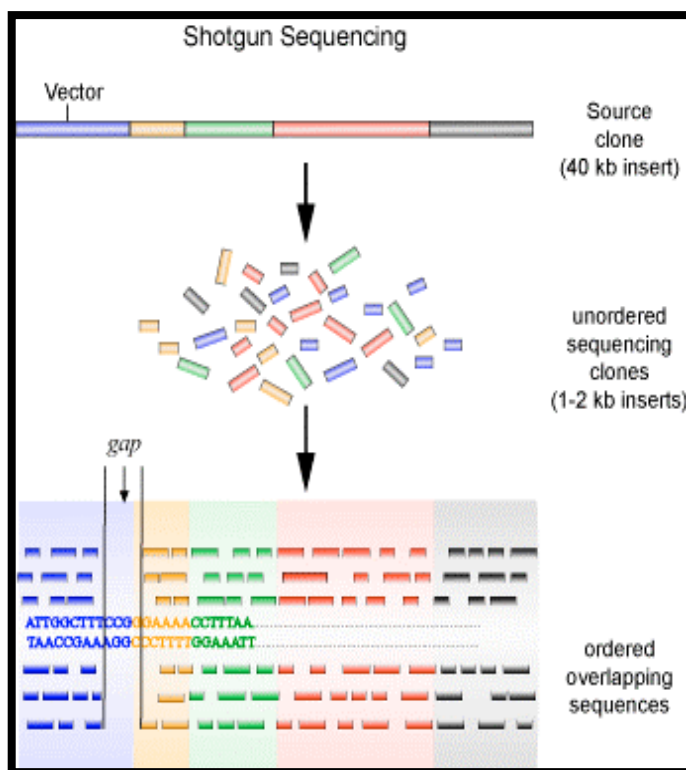
Et av målene i denne oppgaven var å isolere *groESL1* fra *C. aurantiacus*. Ved isolering av spesifikke gener eller genområder er det vanlig å gjennomføre såkalte biblioteker for det ønskede genet. Et bibliotek kan lages ved å isolere organismens kromosomale DNA, kutte DNA'et med et passende restriksjonsenzym og sette de fragmentene man da får inn i en passende vektor. Bakteriofag λ og plasmider er vektorer som ofte brukes og som har den store fordel at de kan oppformeres i en vertsorganisme. Et ideelt genomisk biblioteket inneholder en samling av innsatte fragmenter som til sammen utgjør hele organismens genom. Rekombinante kloner av bakterieceller som inneholder et ønsket gen kan identifiseres på flere måter. Hybridisering er en vanlig metode som innebærer fiksering av denaturert, enkelttrådet DNA fra plakk eller koloni til membran. Membranen dekkes deretter med en væske som inneholder merkede, radioaktivt eller fluorescerende, enkelttrådig DNA eller RNA prober. Membranen og probene inkuberes under forhold som fremmer binding av de merkede probene til komplementært DNA på membranen. Deretter vaskes membranen slik at bare prober som er spesifikt bundet fortsatt vil sitte fast. Hvor proben eventuelt har festet seg kan gjøres synlig ved autoradiografi eller fluorescens, og den plakken eller kolonien som inneholder det ønskede genet kan dermed identifiseres. Å isolere gener fra et bibliotek ved bruk av hybridiseringsteknikker involverer ofte flere runder med søk i biblioteket ved hjelp av utsåing på agarskåler og filterhybridisering. Disse metodene er tidkrevende, ofte kostbart på grunn av radioaktive isotoper og filtre og tilbøyelige til å gi falske positive. Isolering av gener ved bruk av PCR kan være et godt alternativ fremfor å bruke hybridiseringsmetoden. Bruk av PCR forutsetter imidlertid at en del av sekvensen til det ønskede genet er kjent slik at spesifikke primere for genet kan lages. I denne oppgaven ble en metode utviklet av King, 1997 og Yu og Bloem, 1996, brukt til identifisering av *groESL1*. Metoden ble brukt på et Bakteriofag λ basert bibliotek som inneholdt fragmenter fra *C. aurantiacus* kromosomale DNA. Metoden er nærmere beskrevet i delen om metoder i denne oppgaven (del 3.6).

Et av problemene i denne oppgaven var at genomet til *C. aurantiacus* ennå ikke var sekvensert og at 3' ende til *groESL1* ikke var kjent da oppgaven ble påbegynt. Det var derfor ikke mulig å utforme et sett med primere som ville oppformere hele operonet. Dette utgjorde den første utfordring i arbeidet.

1.3.1 Sekvensering av sekundære strukturer i DNA

Isolering av gener hos bakterier vil stadig bli lettere etter hvert som flere genomer blir sekvensert. Gener fra sekvenserte genomer kan klones ved å direkte amplifisere de ønskede genene med PCR og deretter sette dem inn i en passende vektor.

Så godt som all DNA-sekvensering som utføres i dag blir foretatt ved hjelp av Sanger-metoden, en enzymatisk basert metode som benytter seg av spesielt modifiserte nukleotider, såkalte dideoksynukleotider. Flere firmaer utfører sekvenseringsarbeid, blant annet DGI (Doe Joint Genome Institute) og TIGR (The Institute for Genomic Research). Etterhvert som de blir ferdigstilt, blir sekvensene lagt ut i ulike databaser, annotert eller ikke. Sekvenseringen av genomet til *C. aurantiacus* ble påbegynt høsten 2000. I oktober 2004 var *C. aurantiacus* sekvensert ved "shot-gun" metoden og lagt ut på NCBI (figur 1.8). Prosjektet ble lagt ut med accession nummer NZ_AAAH0100000, og besto av sekvensene NZ_AAAH0100001-NZ_AAAH01001142. Proteinkodende gener ble identifisert ved å bruke programmene GeneMark og GeneMarkHMM. Funksjonell annotering var basert på CDD (Conserved Domain Database) og COG (Clusters of Orthologous Groups). DNA sekvenser og de respektive utledete proteinsekvensene er tilgjengelige på BLAST.



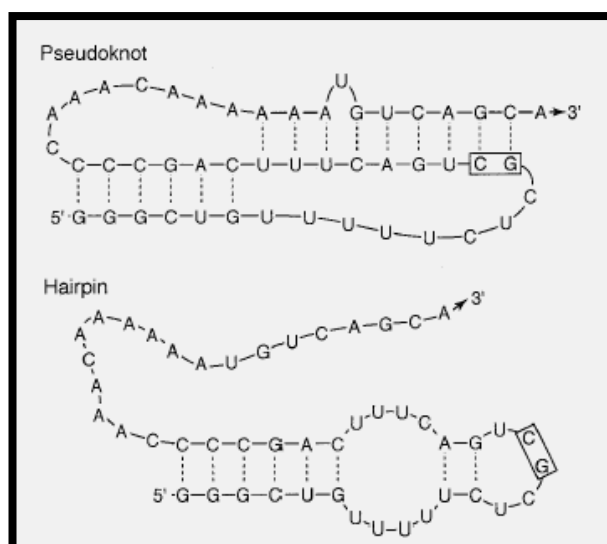
Figur 1.8. Fremstilling av shot-gun sekvensering. I noen tilfeller kan biter av genomet ikke komme med. Dette kalles gaps og er vist på tegningen.

Fordi DNA hos de aller fleste organismer er dobbeltrådet mens RNA er enkelttrådet er RNA mer fleksibelt enn DNA og kan danne flere og mer komplekse tredimensjonale strukturer enn DNA. *In vitro* forsøk, og komputerbaserte forsøk, har vist at enkelttrådet DNA også kan danne komplekse strukturer på lik linje med RNA.

Den lineære sekvensen til DNA og RNA, baserekkefølgen, danner den primære strukturen til nukleinsyrene. Når denne sekvensen foldes, dannes den sekundære strukturen. Kreftene som holder den sekundære strukturen sammen er hydrogenbindinger mellom baseparrene. I RNA eller enkelttrådet DNA kan basene enten være frie (ikke bundet til andre baser) eller bundet til andre baser via hydrogenbindinger. Tertiærstrukturen er den tredimensjonale geometrien av basene arrangert i rommet. Den vanligste hydrogenbindingen mellom basepar som vil føre til dannelse av sekundærstruktur er mellom G og C og mellom A og T/U. Disse bindingene kalles Watson-Crick (W-C) binding. G-C-binding er sterkere da disse basene holdes sammen av tre hydrogenbindinger i forhold til A-T/U bindingene som kun involverer to bindinger. Enkeltrådig DNA og RNA kan foldes i spesifikke tredimensjonale strukturer som bestemmes av nukleotidsekvensen. Forståelsen av RNA-folding er viktig i forbindelse med å forstå en rekke biologiske prosesser, f. eks. translasjonsregulering i mRNA og replikering av enkelttrådig RNA virus. Kunnskap til DNA-folding er viktig blant annet i PCR, der DNA er i enkelttrådet form.

Det er mulig å forutse foldingsstrukturene til enkelttrådet DNA ved hjelp av dataprogrammer utformet til dette formålet (Dirks, R.M., Pierce, N.A., 3003). Mfold programmet baseres på "nearest neighbor" termodynamiske regler for å anslå sekundærstruktur i både RNA og DNA. "Nearest neighbor" regelen går ut på at fri energi er assigned i løkker istedenfor basepar, og kalles derfor også ofte for løkke-dependent regler. Foldingstemperaturen holdes på 37 °C, men kan varieres. Mange av foldingsprogrammene baseres på forandringen i fri energi, ΔG° .

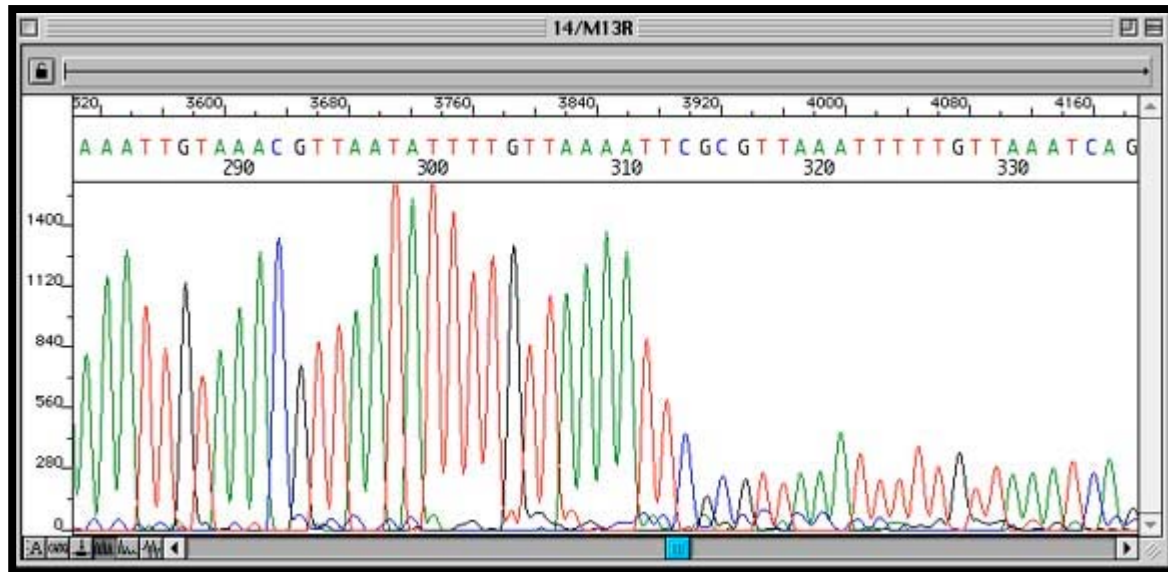
Det finnes visse strukturer enkelttrådet DNA og RNA kan innta og som er kjent. Disse strukturene kalles de elementære strukturene og består av blant annet hårnålsstrukturer og pseudoknuter, som enkle sekundære strukturer eller sammensatt i multidomenestrukturer. En hårnålstruktur opptrer når enkelttrådig område av DNA eller RNA foldes tilbake på tråden, bindes ved hjelp av hydrogenbindinger og danner en løkke. En pseudoknute er en tredimensjonal struktur som oppstår når nukleotidene i en hårnålstruktur bindes til nukleotider et annet sted i sekvensen (figur 1.9). På denne måten blir tre nukleotidtråder holdt sammen av hydrogenbindinger. Hvis deler av sekvensen i en tråd er komplementære, palindrome, kan slike strukturer dannes. Strukturene spiller en viktig rolle i blant annet prokaryotenes transkripsjonsterminering. En sekvens sies å ha sekundærstruktur dersom det finnes et bundet basepar. Dersom sekvensen er helt fri for bindinger mellom basene heter det at sekvensen er strukturfri. De elementære strukturene er lagt inn i dataprogrammer som er laget for å forutsi sekundærstruktur. I slike programmer tas det ikke hensyn til pseudoknuter. Pseudoknuter er heller ikke en av de elementære strukturene fordi det er meget vanskelig å identifisere og dermed forutsi.



Figur 1.9. Bildet viser en pseudoknute og en hårnålstruktur.

Tertiære strukturer mellom allerede bundete baser, f. eks. hydrogenbindinger mellom basene og nukleinsyrenes "rygg", interaksjoner "rygg" mot "rygg", bindin til metallioner, interaksjoner med vann, er fullt mulig og kan virke stabiliserende på molekylerne, men er ikke en del av beregningen av sekundærstruktur. Andre ikke-Watson-Crick bindinger som for eksempel Hoogsteen bindinger virker også inn på den sekundære strukturen. Heller ikke disse formene for sekundærstruktur er med i beregningen, selv om de kan forårsake komplikasjoner under sekvensering.

Områder med høyt G/C-innhold eller sekundære strukturer kan være vanskelige å sekvensere. En hårnålstruktur i DNA templatet vil kunne blokkere for polymerasen i PCR- og sekvenseringsreaksjoner, og på denne måten forårsake at reaksjonen stopper for tidlig. På sekvenseringsgeler vil dette vise seg som like sterke bånd i alle fire sekvenseringsreaksjoner etterfulgt av svake eller manglende bånd. På et kromatogram vil det vise seg som en for tidlig terminering i sekvenseringssignalet. Sekvenseringssignalet, som kan sees som topper for de ulike basene, vil brått ta slutt før sekvenseringen avsluttes (figur 1.10). Årsaken til dette er at DNA-trådene av ulik lengde klumper seg sammen, såkalt komprimering. Komprimeringen kan hindres ved å hindre sekundærstrukturer i DNA-molekylet, eller ved å styrke de denaturerende betingelsene under elektroforesen. Det finnes to analoge former av dGTP som bare danner to hydrogenbindinger med cytosin, og som derved hindrer dannelsen av hårnålstrukturer i G/C-rike enkelttrådede DNA-molekyler. Under sekvensering kan dGTP erstattes av enten 7-deaza-dGTP eller av deoxyinosintrifosfat (dITP). Det er også mulig å hindre dannelsen av sekundærstrukturer ved å øke temperaturen under sekvensering. I visse tilfeller kan det hjelpe å sekvensere den komplementære DNA tråden. Lange A eller T regioner kan også skape problemer under sekvenseringen i form av glipp. Mekanismen bak glipp er enda ikke kjent, men en hypotese går ut på at DNAet ikke holdes riktig parret gjennom polymeriseringen.



Figur 1.10. Sekundærstrukturer i enkelttrådet DNA kan skape problemer i forbindelse med sekvensering av sekvensen. På kromatogrammet vil dette vise seg som en for tidlig terminering i sekvenseringssignalet. Sekvenseringssignalet, som kan sees som topper for de ulike basene, vil brått ta slutt før sekvenseringen avsluttes. Sekundærstrukturen i templatet gir ulike avledninger i sekvensen, noe som kan resultere i ineffektiv elongering.

Der finnes en rekke computerverktøy for biologer, både Windows baserte og UNIX baserte. Et av de største er National Center for Biotechnology Information, NCBI. Her er det informasjon om nukleotidsekvenser og proteinsekvenser, programmer for behandling av slike sekvenser, domenstrukturer, sekvenssammenstillinger, artikler, publikasjoner osv. NCBI har per januar 2005, 331 Bacteria, 23 Archaea og 61 Eukarya genomer inne i BLAST databasene. Disse foreligger i 3 former, avhengig av stadiet sekvenseringsprosjektet har kommet til: ferdige, shot-gun sekvensert og uferdige. Da denne hovedoppgaven ble påbegynt var *C. aurantiacus* enda ikke ferdig sekvensert. Foreløpig ligger den ute som "shot-gun" sekvenser. Mye av arbeidet gjenstår med å finne de overlappende sekvensene og ordne bitene i den rekkefølgen de er i organismens genom. Etter dette følger arbeidet med annotering, finne ut funksjonen til sekvensene, hvilke gener de koder for, om de koder for gener og videre arbeidet med genproduktene.

I denne oppgaven ble ulike programmer brukt som hjelpemiddel til videre arbeid, verifisering av arbeid som er gjort og til innsamling av informasjon. Programmene som ble brukt og deres betydning i arbeidet er beskrevet i oppgavens del om materialer og metoder.

2 MATERIALER

2.1 Kjemikalier

Alle kjemikaliene som ble brukt var av *pro-analysis* kvalitet.

2.1.1 Uorganiske kjemikalier

Forbindelser:	Produsent:
CaCl ₂ x 2 H ₂ O	Merck
HCl, konsentrert	Merck
H ₃ PO ₄	Merck
H ₂ SO ₄	Merck
KH ₂ PO ₄	Merck
K ₂ HPO ₄	NDM
KCl	Merck
MgCl ₂ x 6 H ₂ O	Merck
MgSO ₄ x 7 H ₂ O	Merck
NaCl	AnalaR
NaOH	Merck
NaHCO ₃	Merck
(NH ₄) ₂ S ₂ O ₈ (APS)	Bio-Rad

2.1.2 Organiske kjemikalier

Forbindelser:	Produsent:
Acrylamide	Saveen
Agarose	Amresco
Ampicillin	Sigma
Bacto Agar	Difco
Bacto Trypton	Difco
Bacto Pepton	Difco
Bromfenolblått	Sigma
BSA	Difco
Casamiosyrer	Difco
Coomassie Blue R-250	Applichem
Coomassie Brilliant Blue G-250	Sigma
Dimetyl sulfoksid (DMSO)	Sigma
Dinatriumsalt-dihydrat (EDTA)	AppliChem
Eddiksyre, konsentrert	Merck
Etanol 96%	Arcus
Etidiumbromid	Merck
Fenol	Appligene oncor
Ficoll	Sigma
Gelatin	Mahson Colm & Bell
Gjærekstrakt	Difco
Glukose monohydrat	Merck
Glyserol	Sigma
Isopropylthio-β-D-galaktosid (IPTG)	Calbiochem
Kloroform	Merck
Kloramfenikol	Sigma
NZ-amine	Sigma
Natrium acetat	Merck
Orange G	Merck

Natrium laurylsulfat (SDS)	Merck
Sukrose	NMD
TEMED	Promega
Tetrasyklin	Sigma
Tris-base	Angus
X-gal (X-gal 5-Bromo-4 Chloro-3-indolyl- β -D-galactosid)	Melford Laboratories Ltd

2.2 Primere

Primerene ble utformet manuelt og ved hjelp av det Windows baserte programmet Primer3. Primerene For1, For2, Rew3, GroEL1F og GroEL1R ble utformet av Marc Anglès d'Auriac. Primere ble bestilt av MWG, Genomic Company og ble levert i en konsentrasjon på 100 pmol/ μ l. Eventuelle haler på primerene er merket med uthevet skrift. De vektorspesifikke primerene er standard primere.

Vektorspesifikke primere:

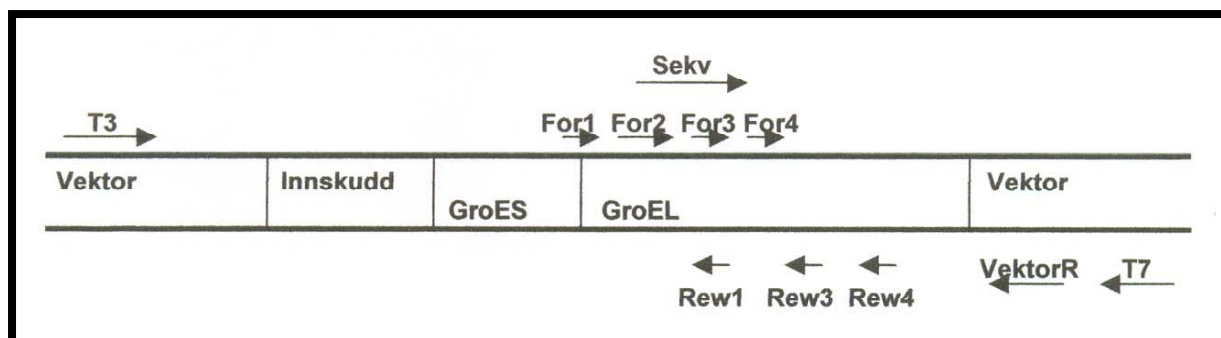
Navn:	Sekvens:	Verdier:
T3	5'-AATTAACCCTCACTAAAGGG-3'	(20bp, tm53.2°C, 40%G/C)
T7	5'-TAATACGACTCACTATAGGG-3'	(20bp, tm53.2°C, 40%G/C)
T7 reverse	5'-TAGTTATTGCTCAGCGGTGG-3'	(20bp, tm58°C, 50%G/C)

Genspesifikke primere til søk etter *groESL1* i bibliotek:

Navn:	Sekvens:	Verdier:
For1	5'-TAAGCAGTATGGCGAAGCAA-3'	(20bp, tm56°C, 45%G/C)
For2	5'-GCGCTGAAACACGGTATTGATA-3'	(22bp, tm69°C, 45%G/C)
Rew1	5'-GCCGAGATGGTAGCGACATTC-3'	(21bp, tm63°C, 57%G/C)
For3	5'-GGTCAAGGCCGACAAAGATA-3'	(20bp, tm57.3°C, 50%G/C)
Rew3	5'-AACCTTAATGACGGCGACAC-3'	(29bp, tm57.3°C, 50%G/C)
For4	5'-GTTTGAAGAAGAGCGGATGG-3'	(20bp, tm57.3°C, 50%G/C)
Rew4	5'-TCACATCGTAGCCGTAGTGG-3'	(20bp, tm59.4°C, 55%G/C)
Gro2	5'-TGGCGGGATGGACTTCTAA-3'	(19bp, tm58°C, 53%G/C)
GATC	5'-AGCCTTCGTTGATGATGG-3'	(18bp, tm54°C, 50%G/C)
Sekv:	5'-TCAATGCCTGCCAGTATAAGCGCCGGGGGG-3'	(30bp, tm77°C, 63%G/C)
VektorR	5'-GCGGCCGCTCTAGAAGTA-3'	(18bp, tm53°C, 61%G/C)

Primere til isolering av *groESL*:

Navn:	Sekvens:	Verdier:
ES1forTOPO	5'- CACCAT GAGGAGGTCGTCTGT-3'	(21bp, tm63°C, 57%G/C)
ES1rew TOPO	5'-TATCAATACCGTGTTCAGCGC-3'	(21bp, tm60°C, 45%G/C)
GroES1F	5'-ATGAGGAGGTCGTCTGTGC-3'	(19bp, tm67.3°C, 58%G/C)
GroES1R	5'-AAAGATCAGTTGCTTCGCCAT-3'	(21bp, tm69.5°C, 43%G/C)
GroEL1F	5'-ATGGCGAAGCAACTGATCTTT-3'	(21bp, tm69.5°C, 43%G/C)
GroEL1R	5'-TGTGCCTACCGGATTGGATTA-3'	(21bp, tm70.9°C, 48%GC)
EL1forTOPO	5'- CACCAT GAGGAGGTCGTCTGT-3'	(21bp, tm61°C, 52%G/C)



Bildet viser en skjematisk fremstilling av noen av primerenes plassering i forhold til hverandre, til vektoren og til innskuddet.

2.3 Standarder

Standarder:	Produsent:
Gene Ruler™ 100bp DNA Ladder	MBI Fermentas
Gene Ruler™ 100bp DNA Ladder Pluss	MBI Fermentas
1kb DNA ladder	Biolabs
Lambda DNA/ <i>Hind</i> III marker2	Fermentas
Predision Plus Protein™ Standards Unstained	Bio-Rad

2.4 Enzymer

Enzym:	Produsent:
Lysozym	Sigma
Proteinase K	Sigma
Rnase A	Sigma
DyNAzyme™ II DNA Polymerase	Finnzymes
DyNAzyme™ EXT DNA Polymerase	Finnzymes
VENT DNA Polymerase	New England BioLabs
Pfu Turbo DNA Polymerase	Invitrogene

I reaksjonene ble polymerasens respektive buffer og eventuelle MgSO₄ brukt som anbefalt av produsentene.

2.5 Utstyr

Standard laboratorieutstyr ble brukt

Utstyr:	Produsent:
Biometra® UNO II Thermocycler	Biometra
Techne Unit Genius Thermocycler	Techne
Mastercycler gradient	Eppendorf
Sonicator	MSE
Pipetter, 0,5-1000µl	Labsystems og Eppendorf
Finntip 65 5ml spisser	Labsystems
1-200µl costar 4783 spisser	Sigma/Heidenreich
200-1000µl Rainin GPS1000 spisser	Tamro Medlab

2.6 Bakteriestammer og bakteriofag

- *Chloroflexus aurantiacus* J-10
- *Escherichia coli* vertsceller til λ ZAP Express™: XL1-Blue MRF⁻, genotyp: $\Delta(mrcA)183 \Delta(mrcCB-hsdSMR-mrr) 173 \text{ endA1supE44 thi-1 recA1 gyrA96 relA1 lac [F}^{\prime} \text{ proAB lac}^q\text{Z } \Delta\text{M15Tn10(Tet}^r\text{)]}$
- One Shot® Top10 Chemically Competent Cells, genotyp: $F^{\prime}mrc \Delta(mrr-hsdRMS-mcrBC)\phi 80lacZ\Delta M15\Delta lacX74recA1deoRaraD139\Delta(ara-leu)7697galUgalKpsL(str^R)\text{endA1nupG}$
- BL21Star™(DE3) One Shot® Chemically Competent *E. coli* B Cells, genotyp: $F^{\prime}ompThsdSB(r_B^{-}m_B^{-})galdcmrne131(DE3)$
- BL21 CodonPlus® (DE3)-RP Chemically Competent *E. coli* B Cells, genotyp: $F^{\prime}ompThsdS(r_B^{-}m_B^{-})dcm+Tet^r gal\lambda(DE3)\text{endA}Hte[argUproLCam]$
- λ ZAP Express™

2.7 Ferdige sett

Champion™ pET 102 Directional TOPO® Expression Kits	Invitrogene
Champion™ pET 101 Directional TOPO® Expression Kits	Invitrogene
QIAquick PCR Purification	Qiagen
QIAprep Spin Miniprep	Qiagen
Wizard® Plus SV Minipreps DNA Purification System Kit	Stratagene
Colony/Plaque Screen™ for Non-Isotopic Detection	NEN™
GFX™ PCR DNA and Gel Band Purification Kit	Amersham Bioscience
S.N.A.P.™ Gel Purification Kit	Invitrogen

2.8 Dataprogrammer:

DNA sekvensdatabaser:

- NCBI, National Center for Biotechnology Information, <http://www.ncbi.nlm.nih.gov/>
- EMBEL-EBI, European Molecular Biology Laboratory-European Bioinformatics Institute, <http://www.ebi.ac.uk/>

Protein databaser:

- Swiss-prot, <http://www.psc.edu/general/software/packages/swiss/swiss.html>
- ExPASy, Expert Protein Analysis System <http://au.expasy.org/>

Primerdesign:

- Primer 3, http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi
- MWG primerdesign, <http://www.mwg-biotech.com/>

Sekvenssammenstillinger:

- Multalin, <http://prodes.toulouse.inra.fr/multalin/multalin.html>
- BLAST, <http://www.ncbi.nlm.nih.gov/BLAST/>
- GenDoc, <http://gendiapo.sourceforge.net/>
- Paralign, <http://www.paralign.org/>

Verktøy

- ORF-Open readingframe Finder, <http://www.ncbi.nlm.nih.gov/projects/gorf/>

- SequenceMassager, <http://www.attotron.com/cybertory/analysis/seqMassager.htm>
- JustBio, <http://www.justbio.com>
- Webcutter 2.0, <http://rna.lundberg.gu.se/cutter2/>
- NEBcutter, <http://tools.neb.com/NEBcutter2/index.php>
- Oligonucleotide Properties Calculator, <http://www.basic.nwu.edu/biotools/oligocalc.html>
- mfold, <http://bioweb.pasteur.fr/seqanal/interfaces/mfold-simple.html>
- SHOWFEAT, EMBOSS, <http://bioweb.pasteur.fr/seqanal/interfaces/showfeat.html>
- Oligo calculator, <http://micro.nwfsc.noaa.gov/protocols/oligoTmcalc.html>
- PRETTYSEQ, <http://cbi.labri.fr/outils/Pise/prettyseq.html>
- PeptideMass, <http://au.expasy.org/tools/peptide-mass.html>
- Rare Codon Calculator, www.doe-mbi.ucla.edu/Services/RACC/ - 4k

Sekvenseringsfirmaer og deres hjemmesider:

- GATC biotech, <http://www.gatc.de>
- MWG The Genomic Company, <http://www.mwg-biotech.com>

Alle programmene som ble brukt var Windows baserte programmer .

2.9 Dyrkningsmedier:

2.9.1 Dyrkningsmedium for *Chloroflexus aurantiacus*:

Rikt medium:

Gjærekstrakt	1.0 g
Casaminosyrer	2.0 g
D-stock*	50 ml
1 M Tris-HCl, pH 8.0	10 ml
Destillert vann til 1 liter	
pH ble justert til 8.2 med 1 M NaOH før autoklaving.	

*D-stock:

Sporelementløsning**	20 ml
Nitriloeddiksyre	2.0 g
CaCl ₂ x 2 H ₂ O	1.0 g
MgSO ₄ x 7 H ₂ O	2.0 g
NaCl	0.16 g
K ₂ HPO ₄ 2.	35 g
Destillert vann til 1 liter.	

**Sporelementløsning:

H ₂ SO ₄ 1M	0.5 ml
MnSO ₄ x 7 H ₂ O	1.14 g
ZnSO ₄ x 7 H ₂ O	0.5 g
H ₃ PO ₄	0.5 g
NaMoO ₄ x 6H ₂ O	0.025 g
CuSO ₄ x 5 H ₂ O	0.025 g
CoCl ₂ x 6 H ₂ O	0.045 g
Destillert vann til 1 liter.	

1 M Tris-HCl pH 8.0:

Tris-base	12.1 g
dH ₂ O	50 ml
pH justeres til 8,0 med konsentrert HCl. Destillert vann til 100 ml.	

2.9.2 Dyrkningsmedium for *E. coli*.

LB-medium:

Bacto-tryptone	10.0 g
Bacto gjærekstrakt	5.0 g
NaCl	10.0 g

Destillert vann til 1 liter.

pH ble justert til 7.5 med 1 M NaOH før autoklaving.

NZCYM medium:

NZ-amin	10.0 g
NaCl	5.0 g
Gjærekstrakt	5.0 g
Casaminosyrer	1.0 g
MgSO ₄ x 7 H ₂ O	2.0 g

Destilert vann til 1 liter.

pH ble justert til 7.5 med 1 M NaOH før autoklaving.

S.O.C.-medium:

Gjærekstrakt	0,5 %
Trypton	2,0 %
NaCl	10 mM
KCl	2,5 mM
MgCl ₂	10 mM
MgSO ₄	20 mM
Glukose	20 mM

Gjærekstrakt, trypton, NaCl og KCl tilsettes destillert vann og pH justeres til 7,0 før autoklaving. Etter avkjøling sterilfjernes MgCl₂, MgSO₄ og glukose og tilsettes.

LB agar:

LB medium

Agar 1.5 % (vekt/volum)

Løsningen ble autoklavert, avkjølt til 50 °C og deretter helt ut i petriskåler.

LB toppagar:

LB medium

Agar/agarose 0,7 % (vekt/volum)

Løsningen ble kokt for å løse opp agarosen og deretter fordelt på glass reagensrør med lokk, 3 ml per rør. Rørene ble deretter autoklavert.

2.9.3 Tilsetninger i dyrkningsmedier

Der antibiotika skulle tilsettes mediet, ble dette tilsatt sterilt etter autoklaving.

Ampicillin stock: (25 mg/ml)

250 mg ampicillin løses i 10 ml dH₂O

Sterilfjernes og lagres på -20 °C

For arbeid med plasmider: bruk 100 µg/ml i LB medium

For arbeid med cosmider: bruk 200 µg/ml i LB medium

Tetrasyklin stock

5 mg/ml løsning i 100 % EtOH

Lagres ved -20 °C

10 µg/ml i LB medium

Kloramfenikol stock:

500 mg kloramfenikol løses i 10 ml dH₂O

Sterilfiltreres og lagres på -20 °C

50 µg/ml i LB medium

40 % glukose (1 liter)

600 ml dH₂O varmes

400g glukose has oppi, litt av gangen

Når glukosen er helt løst, ha oppi resten av vannet

Autoklaveres

X-gal

100 mg X-gal løses i 5 ml dimetylformamid

Oppbevares på -20 °C i eppendorfrør. Rørene må ha folie rundt for å forhindre at lys ødelegger. Det er ikke nødvendig å steril filtrere X-gal løsninger.

2.10 Buffere og løsninger**2.10.1 Buffer for oppbevaring av fag:**SM buffer

NaCl	5.8 g
MgSO ₄ x 7 H ₂ O	2.0 g
1 M Tris-HCl pH 7.5	50.0 ml
Gelatin	5.0 ml av 2 %
Destillert vann til 1 liter	

2.10.2 Løsninger for isolering av DNALøsning 1:

EDTA	20 mM
Sukrose	10 % (vekt/volum)
Tris-HCl pH 8.0	50 mM

Løsning 2:

EDTA	93 g
Tris base	0.605 g
NaOH pellets	ca. 10-12 g (nok til å få pH over 8.0 og for å få løst opp EDTA)
n-Lauroyl Sarcosine (natriumsalt)	5 g
pH justeres til 9.5 med konsentrert NaOH. Destillert vann til 500 ml. Løsningen ble sterilfiltrert før bruk.	

2.10.3 Buffer for oppbevaring av DNATE buffer:

10 mM Tris-HCl pH 8.0
1 mM Na ₂ -EDTA

2.10.4 20 mM Kalium-fosfatbuffer pH 7.5

1 M K ₂ HPO ₄	15ml
1 M KH ₂ PO ₄	5 ml
Destillert vann til 1 liter. pH justeres med 20 mM K ₂ HPO ₄ og 20 mM KH ₂ PO ₄ før autoklaving.	

2.10.5 Buffere og løsninger for agarosegel og elektroforese:Agarosegel:

1 x TAE buffer*	50 ml
Agarose	0,5-2 g, etter ønsket agarprosent
Løsningen varmes til all agarosen er løst i mikrobølgeovn.	

*50 x TAE:

Tris base	242 g
Konsentrert eddiksyre	57,1 ml
0.5 M EDTA pH 8.0	100 ml
Destillert vann til 1 liter	

Påføringsbuffer (O-mix):

Ficoll	20 %
Orange G	0.25 %
EDTA	20 med mer

Etidiumbromid:

10 mg/ml ferdiglaget løsning fra Merck.

2.10.6 Løsninger for arbeid med proteiner:Polyacrylamid gel:Kalkulering for X % separerende gel:

Løsning A	x/4 ml
Løsning B	2,5 ml
dH ₂ O	(7,5 - x/4) ml
10%APS	50 µl
TEMED	5 µl (10 µl hvis x < 8 %)

16 % gel:

Løsning A*	4 ml
Løsning B**	2,5 ml
dH ₂ O	3,5 ml
APS	50 µl
TEMED	5 µl

5 % gel:

Løsning A*	3 ml
Løsning C***	0,67 ml
dH ₂ O	2,3 ml
APS (1)	30 µl
TEMED	5 µl

12% gel:

Løsning A*	3 ml
Løsning B**	2,5 ml
dH ₂ O	3,5 ml
APS	50 µl
TEMED	5 µl

Oppskriftene rekker til 2 stk 6 x 8 cm geler.

(1) 10 % Ammonium Persulfat, APS, 1 ml

0,5 g ammonium persulfat

5 ml H₂O

Kan oppbevares i flere måneder i kjøleskap

*Løsning A: Akrylamid Stamløsning

30 % akrylamide 29,2 g

0,8 % bis-akrylamide 0,8 g

destillert vann til 100 ml

Kan oppbevares i kjøleskap i måneder.

En ferdiglaget løsning, 40 % "ready mix", fra Saveen ble brukt.

**Løsning B: 4x Separeringsbuffer, 100ml

2 M Tris-HCL (pH 8,8) 75 ml

10 % SDS 4 ml

dH₂O 21 ml

Kan oppbevares i kjøleskap i måneder.

***Løsning C: 4x Stacking gel-buffer, 100ml:

1 M Tris-HCL (pH 6,8) 50 ml

10% SDS 4 ml

dH₂O 46 ml

Kan oppbevares i kjøleskap i måneder.

Elektroforesebuffer, 1 liter

25 mM Tris-base 3 g

glycine 14,4 g

SDS 1.0 g

dH₂O til 1 liter

pH = 8,3

Kan også lages i 10 % stock løsning, og oppbevares i romtemperatur for lengere tid.

5x prøvebuffer, 10 ml:

1 M Tris-HCL (pH 6,8) 0,6 ml

50 % glycerol 5 ml

10 % SDS 2,0 ml

β-mercaptoetanol 0,5 ml

1 % bromofenol blått 1,0 ml

dH₂O 0,9 ml

Kan lagres ved -20 °C i måneder.

Coomassie fargeløsning, 1 liter:

Coomassie Blue R-250 1,0 g

Methanol 450 ml

Konsentrert eddiksyre 100 ml

dH₂O 450 mlCoomassie avfargingsløsning, 1 liter:

Metanol 100 ml

Konsentrert eddiksyre 100 ml

dH₂O 800 ml

3 METODER

3.1 Sterilteknikk

Vanlige sterilteknikker for mikrobiologisk og molekylærbiologisk arbeid ble brukt. Alle løsninger og nødvendig utstyr ble sterilisert ved autoklaving, 120 °C i 20 minutter. Løsninger som ikke kunne autoklaveres ble sterilfiltrert. Buffere, løsninger og medium er beskrevet i avsnittet om materialer, med mindre annen informasjon er gitt.

3.2 Dyrking av celler

3.2.1 Dyrking og høsting av *Chloroflexus aurantiacus*

Chloroflexus aurantiacus, stamme J-10, ble dyrket i 60 ml flasker med skrukork i rikt medium ved 55 °C i varmeskap. Lyspærer på 25 og 40 W ble brukt som lyskilde. Flaskene med celler ble plassert ca. 20-30 cm fra lyskilden. Hver 3de uke ble celler overført til friskt, nytt medium. Inokulumet var omtrent 10 % av flaskens innhold. For å få anaerobe forhold må flaskene fylles helt opp med medium. Det nye mediumet forvarmes til 55 °C før overføring av celler. Cellene overføres til flaskene ved hjelp av sterile pipetter.

Cellene som ble brukt til DNA isolering, ble dyrket i 500 ml flasker med skrukork, ved samme dyrkningsforhold som for 60 ml flaskene. Inokulumet til disse flaskene var omtrent 5 %. Cellene ble høstet når OD₆₀₀ var omtrent 1.0, etter ca. 4 døgn, ved sentrifugering i 15 min ved 6 000 rpm ved 4 °C. Cellematerialet ble vasket 3 ganger med 20 mM kaliumfosfatbuffer pH 7.5. Cellene ble oppbevart i fryser ved – 20 °C.

3.2.2 Dyrking av *Escherichia coli*

Stamme *XL1-Blue MRF'* er vertsceller for bakteriofag λ ZAP. Vertscellene ble dyrket ved å inokulere en koloni i 50 ml LB eller NZCYM medium tilsatt 0.2 % maltose og 10 mM MgSO₄. Maltose induserer produksjon av maltose transportprotein (*LamB* genet) hos *E. coli*. Dette proteinet virker også som en reseptor for bakteriofag λ ZAP. Binding av λ til reseptoren er mer effektiv i nærvær av magnesiumioner.

Cellene ble dyrket på risteinkubator ved 37 °C til OD₆₀₀ var ca. 0.5 (4-6 timer). Overnattekulturer ble ofte brukt. Cellene ble da dyrket ved 30 °C for å forhindre at kulturen ble utvokst og at fag dermed ble festet til døde celler. Skulle cellene oppbevares, ble de sentrifugert ned ved 4500 rpm i 10 min ved 4 °C, tilsatt 15-20 ml MgSO₄ og oppbevart ved 4 °C. For å få best mulig effektivitet ved utsåing av fag ble alltid nylagde, 1-2 døgn gamle, vertsceller brukt. Det ble ikke brukt tetrasyklin i disse kulturene, da det er lett å unngå forurensninger i mediet ved bruk av vanlige sterilteknikker.

XL1-Blue MRF' ble oppbevart på LB-agarskåler tilsatt tetrasyklin (12.5 µg/ml) ved 4 °C. Tetrasyklin virker bakteriostatisk, dvs hemmer proteinsyntesen. Tetrasyklinresistens skyldes at cellene har fått den egenskapen at de kan pumpe ut tetrasyklin fra cellen. *XL1-Blue MRF'* ble strøket ut hver uke på LB-tetrasyklinskåler, inkubert ved 37 °C

overnatt og deretter oppbevart ved 4 °C. For langtidsoppbevaring av stammene ble 850 µl flytende kultur tilsatt 150 µl glyserol og frosset ned ved –80 °C.

Andre stammer av *E. coli* ble dyrket etter samme prinsipp. Volumene av rørene som ble brukt til overnatt kulturrene var alltid 5x så stort som volumet av mediet. Bakterier med antibiotikaresistens ble dyrket på LB agar-skåler/flytende medium med tilsvarende antibiotika. Skålene ble inkubert i varmeskap og kulturer i risteinkubator ved 30 °C eller 37 °C.

3.2.3 Utsåing av bakteriofag λ ZAP

Ved utsåing av bakteriofag λ ZAP ble en 10 x fortynningsrekke laget av fagsuspensjonen. En mengde fagsuspensjon (0,5-15 µl) ble fortynnet i SM buffer til ønsket fortynning. En mengde (5-20 µl) av den ønskede fortynningen ble tilsatt 200 µl vertsceller. Rør med fag og versceller ble blandet forsiktig med pippettering, og deretter inkubert ved 37 °C i 15-20 minutter for at fagpartiklene skulle festes til cellene. Blandingen ble deretter hatt oppi 3 ml smeltet LB/NZCYM-medium, 48 °C. Røret ble blandet raskt på vortexmikser og helt ut på en LB/NZCYM-agar-skål. Skålene ble stående i romtemperatur til toppageren hadde stivnet, før de ble snudd opp ned og satt til inkubering i varmeskap ved 37 °C over natt. Antall plakk kunne da bestemmes og titeret (antall fag/ml) i fagsuspensjonen beregnes.

Der skålene skulle brukes til å utføre plakkloft med membran måtte toppageren lages med agarose istedenfor agar. Like mye agarose settes til som om man skulle bruke agar. Dette vil gjøre toppageren litt sterkere etter at den har stivnet og fører til at toppageren ikke fester seg til membranen og følger med når membranen løftes opp.

3.2.4 Lagring og oppbevaring av celler

Bakteriestammer kan lagres i lang tid ved lave temperaturer (- 20 °C og – 80 °C) i 15-40 % glyserol. Dette gjøres på følgende måte:

- 5 ml LB medium, tilsatt eventuelt ønsket antibiotika, inokuleres med en fersk isolert koloni. Kulturen inkuberes ved 37 °C til kulturen er i sen log fase, som regel 6 timer til overnatt.
- For en stamkultur som skal lagres på – 80 °C, pipetteres 225 µl sterilt 80 % glyserol oppi 1.0 ml av bakteriekulturen. Blandingen får en sluttkonsentrasjon på 15 % glyserol. Blandingen blandes godt og settes på – 80 °C.
- For en stamkultur som skal lagres på – 20 °C, pipetteres ut likt volum 80 % glyserol og bakteriekultur før blandingen blandes godt og settes på – 20 °C. Blandingen vil være en flytende stamkultur. Dersom ikke blandingen blandes godt nok, vil iskrystaller dannes og minske kvaliteten på cellene. Cellene undersøkes etter en uke ved å stryke ut en bakterieløkke på agar.

For å bruke en glyserol stamkultur fra – 80 °C tar man en steril tannpirker og skraper av litt av isen, stryker den ut på en LB skål eller annet egnet medium. Stamkulturen tines ikke for hver gang man tar ut celler da dette vil slite på cellene. Hver fryse-tine runde kan føre til et tap på 50 % i cellenes levedyktighet. For å bruke en – 20 °C stamkultur, pipetteres 50-100 µl ut som inokulum, has i 5 ml egnet medium og inkuberes over natt.

3.2.5 Kompetente celler

En celle som kan ta opp DNA og bli transformert er kompetent. I naturen finnes naturlig kompetente stammer innenfor rikene Bacteria og Archaea. Trolig er denne evnen genetisk bestemt, og involverer bestemte proteiner som deltar i opptak og bearbeidelse av DNA.

Celler kan gjøres kompetente *in vitro* ved behandling av ulike kjemikalier. Ved for eksempel å behandle *E. coli* med CaCl_2 og lav temperatur, kan disse cellene ta opp fremmed DNA i form av plasmider. Hva kompetansen kommer av er foreløpig ikke forstått (Sambrook et al. 2000). Det er antatt at cellene kommer i en tilstand der egenskapene til celleveggen og membranen forandres. Når de CaCl_2 -behandlede cellene utsettes for varmesjokk sammen med plasmid DNA, vil plasmidet bli tatt opp. CaCl_2 -transformasjon har en begrenset effektivitet. Ikke mer enn 1 % av cellene vil ta opp plasmid DNA.

Celler uten plasmid ble strøket ut på LB skåler og satt til inkubering i varmeskap ved 37 °C overnatt. En isolert koloni ble plukket med tannpirker og hatt oppi et 50 ml NUNC rør tilsatt 10 ml flytende LB medium. Røret ble satt i risteinkubator (200rpm) ved 37 °C over natt. 1 ml av overnatt-kulturen ble overført til 100 ml nytt LB medium og satt i risteinkubator ved 37 °C i 2-3 timer, til OD_{600} var rundt 0,5. Cellene ble deretter overført til sterile sentrifugerør, inkubert på is i 20 minutter, før de ble spunnet ned (10 minutter / 2500 rpm / 4 °C). Supernatanten ble kastet og bunnfallet resuspendert i 6 ml iskald 0,1 M CaCl_2 . Blandingen ble inkubert ved 4 °C i 30 minutter. Cellesuspensjonen ble igjen sentrifugert ned ved de samme sentrifugeringsbetingelsene. Supernatanten ble kastet og bunnfallet resuspendert ved pipettering i 1,6 ml iskald 0,1 M CaCl_2 tilsatt 15 % glyserol. Porsjoner på 50 µl ble fordelt på kalde 1,5 ml Eppendorfrør. Rørene ble oppbevart på – 80 °C.

Celler kan også gjøres kompetente ved hjelp av elektroporering. Ved å utsette bakteriene for en elektrisk puls med høy spenning vil det dannes midlertidige porer i membranen slik at DNA fra omgivelsen kan taes opp. Denne metoden vil gjøre rundt 10 % av cellene kompetente til å ta opp fremmed DNA (Sambrook et al. 2000).

3.3 Arbeid med DNA

3.3.1 Isolering av DNA fra hele celler

I prinsippet kan DNA isolering fra hele celler deles inn i 3 hovedtrinn:

- Lysering av celler
- Denaturering og fjerning av proteiner/celleavfall med fenol/kloroform
- Felling av DNA med salt og etanol

For lysering av cellene brukes lysozym som ødelegger celleveggen og SDS som løser fosfolipider og proteinkomponenter i celledmembranen. For å fjerne proteiner og celleavfall ekstraheres de lyserte cellene med fenol/kloroform. Proteiner og celleavfall løses i fenol/kloroform, mens DNA vil være i vannfasen. DNAet kan deretter felles med salt og etanol. Etanol fjerner de polare vannmolekylene som omgir DNA-molekylene, slik at positive ioner kan binde seg til ladete fosfatgrupper på DNA, med den følge at DNA felles ut. Etter første felling blir det utfelte DNAet løst og behandlet med Rnase A og Proteinase K, som fordøyer henholdsvis RNA og proteiner. DNA-molekylene er tett

pakket med proteiner, og for at DNAet skal kunne renses, må det frigjøres fra proteinene. Proteinase inaktiverer i tillegg eventuelle aktive Dnaser i lysatet og forhindrer på denne måten videre nedbryting av DNA. Protein- og RNA rester, samt proteasene blir fjernet ved å ekstrahere med fenol/kloroform. DNAet felles til slutt med salt og etanol og løses opp i en egnet stabiliseringsbuffer. Renset DNA kan lagres i TE-buffer som består av Tris-HCl, EDTA og har god bufferkapasitet. EDTA binder divalente kationer (Mg^{2+} , Ca^{2+}) og hemmer på denne måten aktiviteten til enzymer.

3.3.2 Isolering av kromosomalt DNA fra *Chloroflexus aurantiacus*

C. aurantiacus ble dyrket og høstet som beskrevet i punkt 3.2.1.

1 gram våtvekt celler, oppbevart på is, ble tilsatt 10 ml løsning 1 og resuspendert. Cellene ble overført til 10 ml Falcon plastrør. Lysozym ble tilsatt til en sluttkonsentrasjon på 5 mg/ml, det vil si 50 mg. Rørene ble inkubert ved 37 °C i 90 minutter. Etter inkubering ble 10 ml løsning 2 tilsatt, og rørene plassert på vendemaskin ved romtemperatur 1 time, for å få en homogen blanding.

For å fjerne proteiner og celleavfall ble væsken ekstrahert to ganger med et like stort volum fenol/kloroform. Fenolen ble justert til pH 8.0 med pH maker, som beskrevet i tilhørende bruksanvisning. For å få en god separasjon av fasene ble løsningen sentrifugert ved 4000 rpm i 3 minutter ved 4 °C. Interfasen som inneholdt proteiner og den nederste fasen med celleavfall ble kastet. Vannfasen som lå øverst og inneholdt DNA ble beholdt.

Vannfasen ble videre ekstrahert to ganger med like stort volum kloroform. For å felle DNA ble 2,5 volum 96 % etanol tilsatt. Røret ble satt i kjøleskap – 20 °C i 1 time. Røret ble deretter sentrifugert ved 4000 rpm i 3 minutter ved 4 °C. DNA'et ble spunnet ned som hvitt bunnfall. Bunnfallet ble løst i 4 ml 1 x SSC og tilsatt Rnase A til en sluttkonsentrasjon på 100 µg/ml. Etter at røret hadde stått ved romtemperatur i 30 minutter ble Proteinase K tilsatt til en sluttkonsentrasjon på 100 µg/ml. Røret ble inkubert ved 55 °C over natt.

Løsningen ble ekstrahert to ganger med et like stort volum fenol/kloroform og to ganger med et like stort volum kloroform. DNA ble felt med 2,5 volum 96 % etanol ved – 20 °C i 1 time. Bunnfallet ble løst i 200 µl TE-buffer.

For å unngå gjentatt tining og nedfrysing ble DNAet oppbevart i små porsjoner (10-20 µl) i Eppendorfrør. Rørene ble oppbevart ved – 20 °C. Før nedfrysing ble konsentrasjonen av det rensede DNAet målt.

3.3.3 Måling av DNA-konsentrasjon og undersøkelse av kvaliteten til isolert DNA

(Sambrook et al., 1989)

Ved å sammenligne absorpsjonsspekteret til en prøve med et kjent spekter kan man bestemme renheten til prøven. Aromatiske aminosyrer har et markert absorpsjonsmaksimum nær 280 nm. Proteiner som inneholder disse aminosyrene har derfor sterk absorpsjon ved 280 nm og målinger kan derfor brukes til å bestemme konsentrasjon av protein i løsning. DNA og RNA har en tydelig absorpsjonstopp nær 260 nm, mens polysakkarider normalt ikke har noen tydelig absorpsjonstopp i spektrofotometerets bølgelengdeområde.

Konsentrasjonen av isolert DNA ble bestemt ved spektrofotometrisk måling ved 260 nm. For dobbeltrådig DNA gir en konsentrasjon på 50 µg/ml en A_{260} verdi på 1,0. Konsentrasjonen av DNA kan derfor beregnes ut fra formelen:

$\mu\text{g/ml DNA} = 50\mu\text{g/ml} \times A_{260} \times \text{fortynningsfaktoren}$

For å undersøke renheten til preparatet ble A_{280} målt og forholdet A_{260}/A_{280} beregnet. Rene preparater med DNA har et forholdstall på 1,8. Er preparatet forurenset med protein eller fenol vil verdien være lavere enn 1,8, og nøyaktig bestemmelse av konsentrasjonen vil da ikke være mulig. De spektrofotometriske målingene gjøres i kvartskvetter.

Konsentrasjonen av DNA i løsningen kan også undersøkes ved hjelp av agarose gelelektroforese. Av en mengde DNA ble en mengde applisert i en brønn på gel. I en annen brønn ble en mengde og størrelsesstandard, en kvantitativ stige, applisert. Elektroforese ble utført ved 70 V. Ved å undersøke gelen under UV lys kan konsentrasjonen av DNA i løsningen beregnes ved å sammenligne styrken på båndet med båndene i standarden. Inneholder prøven mye RNA kan dette sees som bånd nederst på gelen. Degradert DNA vil være synlig som smear eller i form av at DNA-båndet ikke er helt avgrenset.

Man kan på forhånd beregne det teoretiske utbyttet DNA i en prøve for å se hvor mye cellemateriale man trenger for å få en viss mengde DNA avhengig av hva DNAet skal brukes til. Av tørrvekt bakterieceller vil ca. 50 % være protein og 3 % være DNA.

3.3.4 DNA til PCR reaksjon direkte på celler

Der PCR skulle utføres direkte på celler måtte cellene behandles først for at de skulle lysere og DNA komme ut i løsningen. Dette kan gjøres på ulike måter avhengig av hvilke celler man arbeider med.

En enkel metoden å isolere DNA fra celler til bruk direkte i PCR er:

1. En koloni plukkes med en steril tannpirker og skylles i 500 ml dH_2O . Blandingen fryses og tines 3 ganger med vortexing innimellom repetisjonene for å bryte celleveggen. Dersom celler fra flytende kultur brukes, må cellene spinn ned ved 1500 rpm i 5 minutter. Cellene resuspenderes i 1ml fosfatbuffer og sentrifugeringen gjentas. Dette gjøres to ganger. Denne prosessen vil vaske bort medium fra cellene. Etter siste sentrifugering, resuspenderes cellene i 20 μl dH_2O . Selv om ikke alt DNAet vil frigjøres fra cellene, vil det med begge metodene være nok til å kunne utføre PCR.
2. Prøvene settes på vannbad, ca. 95°C i 5 minutter. Dette trinnet inaktiverer Dnase molekyler. Hvis disse forblir intakt vil det kunne kutte DNAet i fragmenter og DNAet vil ikke lenger være egnet til PCR.

For noen celler er det tilstrekkelig å programmere inn en såkalt "hot-start" i PCR programmet.

3.3.5 Agarose gelelektroforese

Elektroforese er en teknikk som brukes til å separere og rense makromolekyler, spesielt proteiner og nukleinsyrer som varierer i størrelse, ladning og konformasjon. Det er en av de mest brukte metodene innenfor biokjemi og molekylærbiologi. Når ladete molekyler plasseres i et elektrisk felt vil de bevege seg mot den positive eller negative polen, avhengig av den ladningen de har. I forhold til proteiner som kan ha enten en netto positiv eller negativ ladning, er nukleinsyrer alltid negativ ladet på grunn av fosfatet, og beveger seg derfor alltid mot anoden. Makromolekylene

utsettes for elektroforese i en gel med brønner der prøvene blir påsatt. Gelen dekkes med elektroforesebuffer bestående av ioner, som står for strømmen gjennom væsken, og buffer som holder pH'en konstant.

Agarose er polysakkarider ekstrahert fra alger. En konsentrasjon på 0,5-2 % er vanlig å bruke. Ved å variere konsentrasjonen på agarosen kan man skille fragmenter fra 100 bp til 60 kb fra hverandre. Små DNA fragmenter vil vandre raskere enn store fordi de har mindre friksjonstregghet og derfor beveger seg lettere gjennom porene i gelen. Konformasjonen til DNA påvirker vandringshastigheten, sirkulære supercoilede plasmider vandrer for eksempel raskere enn lineære molekyler av samme størrelse. DNA båndene i gelen kan gjøres synlig ved hjelp av ethidiumbromid. Dette kan enten has i gelen før den stivner, eller man kan farge gelen etter elektroforesen. Det fluorescerende stoffet setter seg mellom basene i DNAet. Gelen må plasseres under en UV lampe for at ethidiumbromiden skal kunne synliggjøres. I denne oppgaven ble ethidiumbromid brukt direkte i gelen.

Størrelsen på ukjente DNA fragmenter kan bestemmes ut fra en størrelsesmarkør, en standard, som inneholder DNA fragmenter av en kjent størrelse. Det finnes flere ulike typer, som brukes avhengig av størrelsen til fragmentene man arbeider med. Det finnes også kvantitative markører som gjør det mulig å beregne mengde DNA i prøven som er satt på gelen.

Ved tillaging av agarosegel blir ønsket mengde agarose tilsatt 50 ml 1 x TAE-buffer. Blandingen kokes i mikrobølgeovn på full styrke til agarosen er løst. Løsningen avkjøles til rundt 60 °C og ethidiumbromid tilsettes, dersom dette er ønskelig. 2,5 µl er tilstrekkelig til 50 ml løsning. Konsentrasjonen av ethidiumbromid skal være på ca. 1,5 µg/ml. Løsningen blir deretter helt ut i støpeformen, og ønsket støpekam settes ned. Kammene danner brønner i gelen hvor DNA prøvene skal appliseres. Det finnes en rekke ulike kammer avhengig av mengde og antall prøver som skal analyseres. Når gelen har størknet, overføres gelen til et elektroforesekar og støpekammen tas ut. Elektroforesekarer skal være fylt med 1 x TAE-buffer.

Prøvene som skal analyseres tilsettes 0,3 x volum påføringsbuffer, her brukt O-miks, og påsettes i brønnene. Standarden blir også tilsatt appliseringsbuffer, 0,2 x volum, og påsettes i en av brønnene. Appliseringsbufferen øker tettheten på prøven slik at DNA synker ned i brønnene. Den gir også prøvene farge slik at det er lettere å holde øye med hvor langt prøvene har vandret i gelen etter at elektroforesen er satt i gang.

Elektroforesen utføres ved 60-90 V. Gelen kan fotograferes under UV-lys.

3.3.6 Isolering av DNA fra agarosegel

For å isolere fragmenter og bånd fra agarosegel ble ulike sett og tilhørende protokoll brukt. De settene som ble brukt og produsentene som distribuerer dem er beskrevet i delen om materialer.

Så mye som mulig av prøven med DNA påsettes gelen og elektroforesen utføres som normalt. Prøvene bør ikke utsettes for høy strømstyrke, da dette kan føre til at båndene blir mer diffuse og utflytende enn nødvendig. Når båndene er separert på gelen, tas gelen ut av elektroforesekarer og plasseres under UV-lys slik at båndene kommer til syne. Båndene som skal isoleres kuttes ut fra gelen med en steril skalpell. Når bånd kuttes ut fra gel under UV-lys, må dette gjøres forholdsvis raskt da UV-lys kan forårsake mutasjoner i DNA-molekylene. Det er viktig å fjerne så mye som mulig av gelen rundt DNAet, da tilstedeværelse av agarose kan gi dårligere utbytte. Gelbiten med DNA legges i et sterilt

mikrosentrifugerør, og veies for at riktig mengde løsninger skal kunne tilsettes etter hvert. 2,5 x volumet 6,6 M natriumjodid tilsettes, og gelbiten smeltes i vannbad ved 45 °C. Deretter tilsetter man bindingsbuffer og overfører løsningen til en filterkolonne med en tilhørende oppsamlingskolonne. Flere trinn med sentrifugering og vaskebuffer fører til at man sitter igjen med rent DNA, som analyseres på gel for å bestemme utbytte og for å kontrollere at riktig bånd er rensset. Prinsippet for isoleringen er at DNA binder seg til en filterkolonne mens andre molekyler renner gjennom. DNA vaskes med etanolholdig vaskebuffer før TE-buffer eller vann tilsettes og rehydrerer DNA slik at det kan elueres fra kolonnen. En mer detaljert beskrivelse er å finne i protokollen til settet.

3.4 Primerutforming og optimalisering av PCR

3.4.1 PCR

PCR, Polymerase Chain Reaction, er en teknikk som brukes til selektiv replikering av en bestemt DNA stykke *in vitro* for å produsere en stor mengde av denne DNA sekvensen.

En PCR reaksjon må optimaliseres for at den skal gi best mulig resultater. Omtrent alle trinnene i reaksjonen kan optimaliseres, men det finnes retningslinjer for hvordan man får frem det mest ønskelige produktet. Optimalisering av en PCR reaksjon har til formål å oppnå det ønskede produkt og øke kvaliteten til dette produktet. Den viktigste delen av optimaliseringen er å finne den riktige balansen mellom dannelsen av det ønskede produktet på den ene siden og dannelse av feilprodukter og primer-dimer på den andre siden. Med en gang en feil-priming har funnet sted vil det resulterende feilproduktet i etterfølgende sykler passe perfekt til primerene, og mengden av det øke samtidig med mengden av det ønskede produktet.

PCR ble prøvd ut med primersett 1 og 2, på genomisk *C. aurentiacus* DNA og på bibliotek, med ulike polymeraser, ulike konsentrasjoner og ulike programmer. Ulike typer PCR maskiner ble brukt, Biometra, Techne og Techne Gradient. Gradientmaskinen ble benyttet for å optimalisere reaksjoner, der det var vanskelig å få frem produkt eller der det var store forskjeller mellom annealingstemperaturene til primerene. Før søk i bibliotek ble utført var det viktig å undersøke om de genspesifikke primerene ville amplifisere ønsket sekvens og dermed også finne ut om biblioteket inneholdt amplifiserbare positive fagklon. PCR produktene ble derfor sendt til sekvensering før primerene ble brukt til søk i bibliotek. Sekvenseringen av korte fragmenter, < 900 bp, ble utført på MegaBace maskinen.

Når PCR ble utført i denne oppgaven ble, det i de tilfellene hvor det ble laget mange prøver med samme reagenser brukt en mastermiks. En mastermiks vil si at alle komponentene til PCR blandes sammen og deretter fordelt på det antall rør som var aktuelt å bruke. Ved pipettering av små volumer kan pipetteringsfeil oppstå. Ved bruk av mastermiks kan større volumer pipetteres av gangen, noe som gjør at sannsynligheten for pipetteringsfeil avtar. Er et av komponentene i reaksjonen ulikt for reaksjonene tilsettes dette separat etter mastermiksen. Polymerasen has oppi helt til slutt. For at ikke komponentene i mastermiksen skal begynne å reagere før de settes i PCR maskinen, er det viktig at dette gjøres på is.

3.4.2 Primerutforming

Design av gode, brukbare primere er en forutsetning for en vellykket PCR reaksjon. Primersekvensene bestemmer produktets lengde, smeltetemperaturen, sekvensens start og slutt og til syvende og sist også mengden produkt.

Primerene må sjekkes mot målsekvensen for å forsikre om at det ikke er andre steder på sekvensen der den kan anneale. Dette kan gjøres manuelt med korte sekvenser eller ved bruk av et av de mange alignmens-programmene på nett. Dersom primersetet skal amplifisere et helt spesifikt område kan ikke primerdesign-programmer brukes. Primerene må da utformes manuelt, men programmer som Oligonucleotide Properties Calculator kan brukes til slutt for å undersøke primerene. Ulike programvarer på nett ble brukt til primerdesign, men ble også gjort manuelt. Listen over de ulike programmene er å finne i oppgavens del om materialer.

3.4.3 Utforming av primere til søk etter positiv klon

Genspesifikke primere ble konstruert ut ifra den kjente sekvensen til *groESL1*. Primerutforming ble utført ved hjelp av Primer 3 programmet og med hjelp fra Marc Anglès d'Auriac. De to sekvensene *groESL1* og 2 fra *C. aurantiacus* ble sammenstilt i programmet Multialign, og primere laget ut fra sekvensene med minst konsensus. For å få primerene mest mulig spesifikke ble de konstruert forholdsvis korte, 20-22 bp, med en AT-rik 5'ende. Primerene ble laget uten selvkomplementaritet og med et GC innhold på rundt 50 % for å gi høyere annealingstemperatur. Primerene ble også kontrollert manuelt.

Det ble utformet to sett med primere, der revers primeren var lik for begge settene, For1, For2 og Rew1. Primersettene ble testet i programmene BLAST og multalin mot genomene til operonet *groESL1*, *groESL2*, Lambda ZAP og *XL1-Blue MRF'* for å teste spesifisiteten og se etter uønskede bindinger. I PCR reaksjonene som skulle utføres var templatet eluat fra skåler med både plakk og celler. Det var derfor viktig å sjekke primerene mot alle genomene. Primersett 1 ville gi et produkt på 460 bp, mens primersett 2 ville gi et produkt på 420 bp. Primerene ble testet i PCR før bruk, både på bibliotek, *XL1-Blue MRF'* og på genomisk DNA fra *C. aurantiacus*. Reaksjonsbetingelsene ble optimalisert med tanke på konsentrasjoner, Tm, polymerase og tid.

Andre primersett som ble brukt til samme formål videre ut i oppgaven ble utformet og behandlet på samme måte som For1, For2 og Rev1.

3.4.4 Sekvenseringsprimer

Det ble gjort forsøk på å få resten av sekvensen til operonet ved bruk av en lang sekvenseringsprimer. Primeren ble utformet fra et område av den allerede kjente sekvensen og ble brukt på kromosomalt DNA i MegaBasemaskinen. De to *groESL* operonene fra *C. aurantiacus* ble sekvenssammenstilt i multalin og en primer utformet fra et område der de to sekvensene var mest mulig ulike. Primeren var 30 bp lang og hadde en høy

annealinstemperatur, 77 °C. Reaksjonsbetingelsene ble beregnet og reaksjonen satt opp. Reaksjonene ble satt opp i samarbeid med Bård Mathisen.

3.4.5 Revers primer utformet fra *groESL2*

En revers primer ble utformet fra slutten av operonet *groESL2*. Sammen med *groESL1* genspesifikk forover primer ble den prøvd på genomisk DNA og mot bibliotek. Chaperoniner er lite konserverte i områdene mot slutten av sekvensen. Alignments med *groESL1* og 2 viste 67 % likhet.

3.4.6 Konsensusprimer

Hovedprinsippet med PCR er knyttet til det å finne primere som flankerer endene av DNA-fragmentet som skal oppformerer. Å finne slike primere krever at sekvensen til DNA-stykket er kjent. Ved å sammenligne sekvensene til gener av samme familie er det kun mulig å utforme primere ut ifra konserverte domener og konsensussekvenser. Det er også mulig å sammenligne aminosyresekvenser og lage degenererte primere ut fra aminosyrenes konsensussekvens.

Chaperoniner er noen av de mest konserverte proteinene vi kjenner. Aminosyresekvensen til *groEL* fra ulike termofile organismer og ikke-termofile organismer ble hentet fra BLAST og sammenlignet for å finne et konservert område der en degenerert primer kunne utformes. GeneDoc programmet ble brukt til dette. Ved å bruke konsensussekvensen til sekvenssammenstillingen skulle en revers primer utformes. Denne primeren skulle brukes sammen med primeren for1 på genomisk for å prøve å få opp hele sekvensen.

3.5 Genbiblioteket

3.5.1 Konstruksjon av biblioteket

Et genbibliotek for *C. aurantiacus* var konstruert av Christin Vanberg (hovedoppgave, UIO 2001). λ ZAP Express™ vektor var brukt til å konstruere biblioteket. λ ZAP Express™ vektor er en innsettingsvektor med 12 unike klonings seter hvor DNA med størrelse på 0-12 kb kan settes inn. Innsetting av DNA i vektoren inaktiverer β -galaktosidase genet og tillater blå/hvit fargeseleksjon på X-gal plater, kloner med og uten innskudd kan dermed skilles fra hverandre. Vektoren inneholder også T7 og T3 promotorene som kan brukes til å lage RNA transkripter i fra det fremmede DNA . Sekvensene til promoterene brukes også som primersett i PCR for å karakterisere og sekvensere innskuddet. Kromosomalt DNA fra *C. aurantiacus* var delvis kuttet med restriksjonsenzymet *Sau3AI* og ligert inn i kuttetsetet *BamHI*, da disse danner kompatible overheng.

Biblioteket var amplifisert som beskrevet i protokollen til ZAP Express Predigested Vector Kit. Biblioteket ble oppbevart ved – 80 °C. Da det ble tatt i bruk ble det oppbevart ved – 4 °C.

3.5.2 Test av biblioteket

Da biblioteket ikke var brukt siden 2001, ble det utført to tester for å se at det fortsatt kunne brukes, at lambda ZAP ville infisere vertscellene, og at de fortsatt inneholdt innskudd. Testene var blå/hvit bakgrunnstest og PCR med vektorspesifikke primere.

3.5.3 Blå/hvit bakgrunnstest

For å se på forholdet mellom rekombinante og ikke-rekombinante fag ble blå/hvit fargeseleksjon utført på biblioteket. Vektorsystemer som inneholder lac-promotor-operator-regionen fra *E.coli*, etterfulgt av et klonings-sete i lacZ, kan brukes i denne testen. Dette genet koder for den N-terminale delen av *E.coli* enzymet β -galaktosidase (figur 3.2). Ekspresjon av lacZ genet er under kontroll av lac-promotoren som i fravær av glukose kan induseres av laktose eller isopropyl-1-thio-B-D-galactopyranosid, IPTG, en ikke-metaboliserbar inducer. LacZ polypeptidet uten C-terminal ende har ingen enzymatisk aktivitet og er meget ustabil i *E.coli*. Når det kobles sammen med den manglende C-terminal delen, dannes en funksjonell β -galaktosidase. Enzymet blir funksjonelt når lacZ uten C-terminal uttrykkes i en vertsstamme som også uttrykker den C-terminale delen av β -galaktosidase.

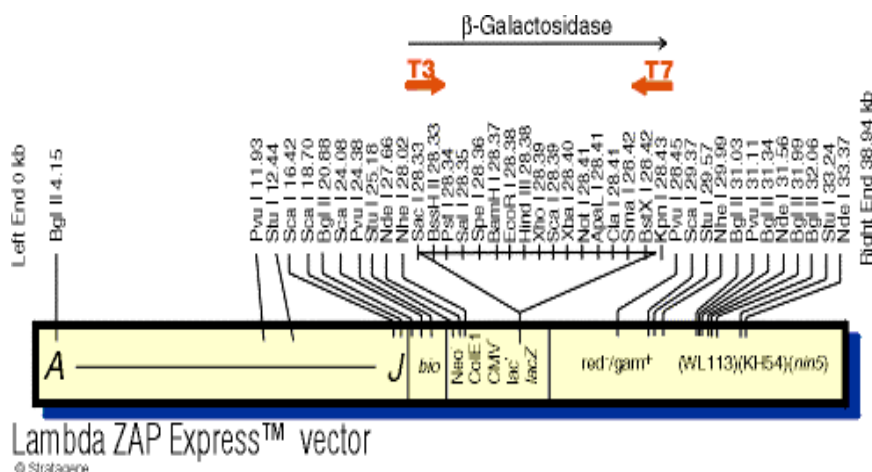
Ved å tilsette IPTG i toppagaren vil lac promotoren i vektoren bli induisert av IPTG og en aktiv β -galaktosidase lages hvis et innskudd er satt inn i vektoren. Ved å inkludere et substrat for β -galaktosidase, X-gal, i toppagaren, kan rekombinante kloner gjenkjennes. X-gal produserer et blått substrat når det blir hydrolysert og følgelig vil ikke-rekombinante plakk være blå. Rekombinante kloner kan ikke lage aktiv β -galaktosidase og plakkene vil derfor være hvite.

Til 3 ml av NZCYM toppagar (48 °C) ble 15 μ l med 0.5 M IPTG og 50 μ l med X-gal (250mg/ml) tilsatt. 100 μ l med *XL1 Blue MRF'*, OD₆₀₀ = 1, ble inkubert med 100 μ l 100x fortynnet bibliotek ved 37 °C i 15 min, før dette ble tilsatt softageren sammen med IPTG og X-gal. Softageren ble helt på LB medium skåler, fikk størkne på benken 20 min, før inkubert over natt ved 37 °C.

Etter at cellene har vokst opp kan forholdet mellom blå og hvite kolonier beregnes og antall rekombinante og ikke-rekombinante fag anslås.

3.5.4 Test av innskudd ved bruk av PCR

Det ble utført en test av bibliotekets innskudd ved hjelp av PCR. Vektorspesifikke primere, T7 og T3, ble brukt for å hente opp innskudd fra biblioteket. Disse primerene binder på hver sin side av kuttesetet til restriksjonsenzymet *BamHI*, som var brukt til å ligere inn fragmentene i vektoren (figur 3.2). T3 virker som en forover primer og T7 som rewers primer. Innskudd i størrelsen 2-5 kb var forventet (Vanberg, Hovedoppgave UIO, 2001) samt noen mindre, degraderte produkter. Bånd på ca. 300 bp ville vise produktet av primersetet på vektorer uten innskudd. PCR produktene ble analysert ved agarose gelelektroforese. Resultatet ville gi en viss formening om størrelsen til innskuddene, mengde fag uten innskudd og om det fortsatt var mulig å hente ut innskuddet ved hjelp av PCR.



Figur 3.2. Bildet viser en tegning av vektoren, med en oversikt over genet for β -galaktosidase, og også plasseringen av vektorspesifikke primere, T3 og T7 (Stratagene)

Reaksjonen ble satt opp:

Reagenser:	Sluttkonsentrasjon:	Volum:
10x VENT polymerasebuffer	1x	5 μ l
10mM dNTPmix	200 μ M	1 μ l
Primer T3	0.4 μ M	2 μ l
Primer T7	0.4 μ M	2 μ l
Bibliotek		2 μ l*
VENT polymerase	0.02 U/ μ l	1 μ l
MgSO ₄	5 mM	5 μ l
dH ₂ O til totalvolum		50 μ l

* 2 μ l var mengden bibliotek som tilsa at hele genomet var representert (Vanberg, Hovedoppgave UIO, 2001).

Det ble satt opp 11 like reaksjoner samt en blindprøve, uten templat. Prøvene ble kjørt i PCR gradient maskin der annealingstemperaturen gikk fra 40 - 60 °C. Tiden for elongering ble på grunn av den forventede størrelsen til produktene satt til 1 minutt og 40 sekunder. Resultatet ble analysert på 1 % agarosegel.

3.6 Søk i bibliotek ved bruk av PCR

I denne oppgaven skulle en positiv fagklon identifiseres ved bruk av PCR. Søket i biblioteket ble gjort etter metoden King (1997) og etter metoden Yu og Bloem (1996). Disse to metodene ble kombinert, og metodene noe modifisert.

Biblioteket var allerede konstruert. Titeret på biblioteket måtte beregnes på nytt, da det var amplifisert. Ut ifra hva man visste om genomet til *C. aurantiacus* og konstruksjonen av biblioteket ble antall fag nødvendige for å representere hele genomet bestemt.

Negativ kontroll ble alltid utført for å undersøke om forurensninger fant sted. Den negative kontrollen ble ikke tilsatt DNA. I tillegg ble en positiv kontroll utført, der templatet var 2 µl bibliotek og primersettet for2/rew1, da dette viste seg alltid å gi bånd. PCR produktene ble analysert ved elektroforese i 1 % agarosegel ved 70 V.

3.6.1 Metoden til King

(King, 1997)

Metoden til King går ut på å snevre inn biblioteket ved bruk av primærsøk, sekundærsøk, tertiærsøk og kvartærsøk. I primærsøket starter man med ca. 5 skåler, hver med 2-3 000 plakk. De positivt identifiserte skålene spres videre ut på skåler som identifiseres gjennom sekundær og tertiærsøket, til man i kvartærsøket sitter igjen med rundt 50-100 plakk. Disse kan isoleres og analyseres separat. Skålene analyseres ved å pipettere SM buffer over skålene. Skålene inkuberes over natt ved 4 °C med SM bufferen. Plakk og celler vil i løpet av denne tiden diffunderer ut i bufferen. Bufferen pipetteres av skålene og sentrifugeres. I supernatanten vil man finne fagen. Dette eluatet kalles for et subbibliotek. Et subbibliotek analyseres med PCR.

Vertscellene som ble brukt i primærsøket ble dyrket opp ved å inokulere en koloni i 50 ml NZCYM medium tilsatt 1 mM MgSO₄ og 0.3 % maltose. Inkuberingen foregikk i rysteinkubator ved 37 °C til OD₆₀₀ var 0.5-1.0. 200 µl celler ble brukt per 100 mm agarskål. Fagene ble sådd ut på NZCYM agarskåler med 3 ml NZCYM toppagar og inkubert overnatt i 12-16 timer.

Primærsøket ble utført ved å så ut omkring 10 000 fag på 5 agarskåler, dvs. omtrent 2 000 fag per skål. Det ble på forhånd beregnet ut ifra titer til biblioteket hvor mye av en fagfortynning det skulle tilsettes til 5 Eppendorfrør med vertsceller. Etter inkubering overnatt ved 37 °C ble 3.5 ml SM buffer pipettert over agarskålene og inkubering foregikk overnatt ved 4 °C. SM bufferen som nå inneholdt fag fra agarskålene ble pipettert av skålene og over i hvert sitt 2 ml mikrosentrifugerør. Celler og agar ble sentrifugert ned ved 4 500 rpm i 10-15 min i en mikrosentrifuge. 1 ml av hver supernatant ble overført til et Eppendorfrør og tilsatt kloroform til 0,3 %, de såkalte subbibliotekene ble oppbevart ved 4 °C. Fra hvert subbibliotek ble 1 µl brukt i PCR. Hvis en eller flere positive fagklon var tilstede på de fem agarskålene ville dette gi et bånd med størrelse tilsvarende 400/460 bp i agarosegelen etter PCR på subbiblioteket. Et positivt subbibliotek var utgangspunktet for fortsettelsen med Yu og Bloems metode. Fra Christin Vanbergs hovedoppgave, 2001, ble det erfart at metoden etter King ikke gjorde det mulig å isolere enkle plakk. Derfor ble Kings metode bare brukt til å påvise primære subbibliotek i denne oppgaven. Herfra ble metoden etter Yu og Bloem brukt.

PCR reaksjonen bestod av:

Reagenser:	Sluttkonsentrasjon:	Volum:
10x VENT polymerasebuffer	1x	5 µl
10mM dNTPmix	200 µM	1 µl
For1 primer	0.2 µM	1 µl
Rew1 primer	0.2 µM	1 µl
Eluat		5-20 µl
Vent polymerase	0.02 U/ µl	1 µl
MgSO ₄ Vent	4mM	2 µl
dH ₂ O til totalvolum		50 µl

PCR programmet som ble brukt:

Program:	Temperatur:	Tid:	Antall sykluser:
Initiell denaturering	94 °C	5 min	x 1
Denaturering	94 °C	30 sek	
Annealing	62 °C	30 sek	x 30
Elongering	72 °C	30 sek	
Siste elongering	94 °C	5 min	x 1
Hold	4 °C	∞	

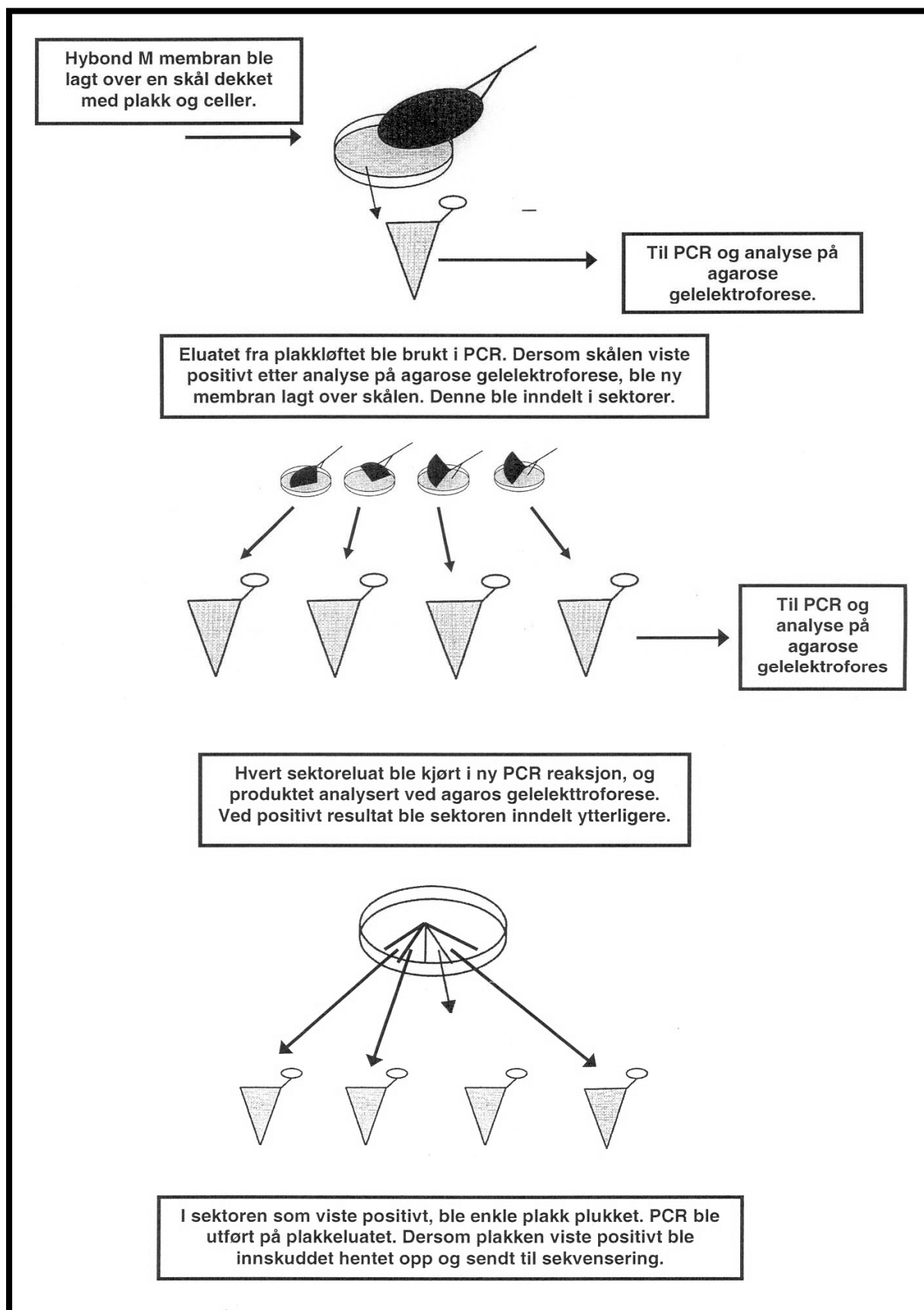
Disse reaksjonsbetingelsene ble også brukt i metoden til Yu og Bloem.

3.6.2 Metoden til Yu og Bloem

(Yu og Bloem, 1996)

Metoden til Yu og Bloem baserer seg på bruk av en innsnevringemetodikk der skåler med fag og celler ikke dyrkes opp på nytt, men heller deles inn i sektorer. Istedenfor sub-biblioteker får man ved bruk av denne metoden såkalte under-seksjoner. Fordelen med denne metoden er at man ikke overfører positivt identifiserte plakk til nye skåler, men beholder dem der de er identifisert. På denne måten vil man kunne unngå å miste spesifikke klon. Ulempen med den er at det krever mange skåler, da man må begynne første søk med godt isolerte plakk, ikke mer enn ca. 200 per skål.

Et av de positive primære sub-bibliotekene, fra primærsøket i metoden King, ble brukt til å så ut 20 skåler med omtrent 200 fag per skål. Etter at fagene hadde dannet plakk ble skålene avkjølt ved 4 °C i 60 min. Følgende prosedyrer ble utført for de 20 skålene: en Hybond N membran ble lagt over agaroverflaten på en skål i ca 10 sek, til den ble fuktet gjennom. Det ble passet på at ikke luftbobler kom mellom membranen og agaren. For å unngå at toppagaren fulgte med ble membranen løftet forsiktig med en steril pinsett og plassert med fagsiden ned i en steril petriskål som inneholdt 3 ml SM buffer. Fagpartiklene ble skylt av membranen ved å løfte membranen opp og ned et par ganger før den ble holdt i 10-20 sek over skålen for at væsken skulle dryppe av. Membranen ble så kastet. 20 µl av SM bufferen med fagpartikler, såkalt fageluut, ble brukt som templat i PCR reaksjonen. Resten av eluatet ble frosset ned, - 20 °C. En skematisk fremstilling av metoden er vist på figur 3.4.



Figur 3.4. Skjematisk fremstilling av metoden etter Yu og Bloem.

10 µl av hvert PCR produkt ble analysert ved elektroforese på 1,8 % agarosegel ved 70 V. Hvis en eller flere positive fagklon var tilstede på de 15 agarskålene ville dette gi et bånd på 460 bp i agarosegelen etter PCR på fageluatet. Den tilhørende agarskålen til et positivt fageluat ble regnet som en positiv skål.

En ny membran ble lagt over den positive skålen. Med en steril skalpell ble membranen og toppagaren under skåret i 4 like store seksjoner. Hver membranseksjon ble løftet av og skylt i hver sin petriskål med 1 ml SM buffer. 20 µl fra hver av de 4 fageluatene ble brukt i PCR. Hvis et bånd på 460 bp var tilstede på agarosegelen ble den tilsvarende seksjonen på agarskålen regnet som positiv.

En ny membran, lik størrelsen på seksjonen, ble lagt over den positive seksjonen på agarskålen. Membranen ble skåret til med en steril skalpell i 4 nye seksjoner. Hver seksjon ble igjen løftet av og skylt i hver sin skål med 0.5 ml SM buffer. 20 µl av fageluat ble brukt i PCR. PCR produktene ble igjen analysert på agarosegel ved elektroforese.

Etter identifisering av en positiv underseksjon i en positiv seksjon ble alle plakkene i den positive underseksjonen stukket ut av agaren med en steril glasspipette og hatt oppi 50 µl SM buffer. For at plakken skulle få diffundere ut i SM bufferen sto rørene ved romtemperatur 2 timer. PCR ble utført med 5 µl som templat. Etter identifisering av den positive plakken, ble denne fagen amplifisert. 10 µl av eluatet ble tilsatt 90 µl SM buffer, blandet med 100 µl XL1-Blue MRF⁺ celler fra en ny ON kultur og satt i varmeskap på 37 °C i 15 min. Deretter ble det ved hjelp av softagar spredt ut på en petriskål med nytt LB medium. Denne skålen ble satt ved 37 °C over natt. Skålen ble avkjølt 1 time på 4 °C, før plakkløft ble utført. Membranen ble skylt i 3 ml SM buffer. Eluatet ble oppbevart ved 4 °C tilsatt kloroform til 3 %.

To ulike metoder ble prøvd med tanke på PCR på isolerte plakk. Den ene metoden, som viste seg å fungere best og som ble brukt videre var metoden der plakken ble plukket med en steril glasspipette. Den andre metoden baserte seg på å stikke en steril tannpirker borti plakken og deretter skylle den direkte i PCR reksjonsblandingen (Dorfman, 1993). Denne siste metoden kan skape problemer når det gjelder å gå tilbake til plakken. Plakken kan merkes på undersiden av agarskålen, men er plakken tett, kan dette være vanskelig. Tannpirkeren ble også skyllet i 20 µl SM buffer, hvorav 10 µl ble brukt videre i PCR og 10 µl til utsåing av ny skål. Denne metoden ga ikke tilfredsstillende resultater.

3.6.3 Isolering av innskudd fra positive fagklon

For å få amplifisere og isolere innskuddet i positive fag ble vektorspesifikke primere, T7 (revers) og T3 (forover), brukt. Dette ble gjort for å undersøke størrelsen av innskuddet og for å få det opp slik at det kunne sendes til sekvensering.

PCR betingelsene var:

Reagenser:	Sluttkonsentrasjon:	Volum:
10x Dynazyme EXT polymerasebuffer	1 x	5 µl
10mM dNTPmix	400 µl	2 µl
T7 primer	0.4 µM	1 µl
T3 primer	0.4 µM	1 µl
Eluat		5 µl
Dynazyme EXT polymerase		1 µl
dH ₂ O til totalvolum		50 µl

Program:	Temperatur:	Tid:	Antall sykluser:
Initiell denaturering	94 °C	5 min	x 1
Denaturering	94 °C	1 min	
Annealing	52 °C	1 min	x 30
Elongering	72 °C	1 min	
Siste elongering	94 °C	10 min	x 1
Hold	4 °C	∞	

PCR produktene ble undersøkt på agarosegel, og fragmenter av ønsket størrelse analysert i ny PCR reaksjon.

3.6.4 Test av innskudd fra positiv fagklon ved hjelp av PCR

Ved å bruke genspesifikke primere i kombinasjon med vektorspesifikke primere ble det mulig å undersøke retningen av innskuddet i vektoren og hvor i fragmentet de genspesifikke primerene ble festet. Størrelsen på PCR produktene ville vise om hele operonet befant seg i innskuddet. Primersettene ville også gi en indikasjon på om fragmentet var verdt å sekvensere. Et kort produkt av T7 og For1 ville være av liten interesse da det var forlengelse av sekvensen i denne retningen det var ønsket mer informasjon.

En annen og enklere måte å undersøke fragmentet på, er å benytte restriksjonskutting. Dette gjøres ved å benytte seg av et restriksjonsenzym som kutter et kjent antall ganger i sekvensen. Ved å analysere kuttingen på gel kan man se om sekvensen er den ønskede sekvensen ut fra antall fragmenter og størrelsen til disse. Slike enzymer kan man lett finne ved å bruke det windows baserte programmet webcutter. Her limer man inn sekvensen og velger ut restriksjonsenzymer. Programmet vil gjøre en kuttereaksjon *in silico*. Dette kunne ikke gjøres i dette forsøket, da sekvensen ikke var kjent.

PCR reaksjonen bestod av:

Reagenser:	Sluttkonsentrasjon	Volum
10x Dynazyme EXT polymerasebuffer	1 x	5 µl
10mM dNTPmix	200 µl	1 µl
for1/T3 primer	0.2 µM	1 µl
T7/rew1 primer	0.2 µM	1 µl
PCR produkt positiv fagklon		1 µl
Dynazyme EXT polymerase		1 µl
dH ₂ O til totalvolum		50 µl

PCR programmet som ble brukt:

Program:	Temperatur:	Tid:	Antall sykluser:
Initiell denaturering	94 °C	5 min	x 1
Denaturering	94 °C	1 min	
Annealing	52 °C	30 sek	x 25
Elongering	72 °C	1 min	
Siste elongering	94 °C	5 min	x 1
Hold	4 °C	∞	

Den lave annealingstemperaturen som ble brukt, 52 °C var et kompromiss mellom primersettene T7/T3 og for1/rew2.

PCR produktene ble analysert på agarosegel.

3.7 Sekvensering av positive fagklon

Fragmenter som viste riktig størrelse og riktig orientering ble sendt til sekvensering. Fordi operonet var antatt å ha en størrelse på 2000 bp, var fragmenter over denne størrelsen av interesse. Flere fragmenter ble sendt til sekvensering. Firmaene som ble benyttet var GATC Biotech og MWG The Genomic Company. GATC biotech i Tyskland, utførte Primer Walking / Single Strand Sequencing på plasmider og PCR produkter med en garantert kvalitetskontroll på >99,9 %. Firmaet MWG The genomic Company ble også brukt. Her ble primer walking servicen benyttet, egnet for produkter på 0,7-4 kb og med en garantert kvalitetskontroll på >99 %.

Ved sekvensering av korte sekvenser ble MegaBACE maskinen benyttet, som er lokalisert på Blindern i Kristine Bonnevis hus. Denne maskinen sekvenserer opp til 800 baser, uten bruk av primerwalking, og ble derfor ikke benyttet ved sekvensering av sekvenser større enn dette.

PCR produkter som skulle sekvenseres på MegaBace maskinen ble levert i eppendorfrør, 5 µl av hver prøve. Prøvene blir oppbevart i kjøleskap for kortere perioder, ved lengre tids lagring settes de på -20 °C.

Da det viste seg å være vanskelig å sekvensere de positive innskuddene, ble det forsøkt å tilsette ulike forsterkende detergenter i PCR reaksjonen. For å øke utbyttet og spesifisiteten i en vanskelig PCR reaksjon kan man tilsette forsterkende virkemidler til reaksjonen, som for eksempel Betaine og DMSO (dimetyl sulfoxide). Av DMSO bruker man en løsning på 2-10%. 10% DMSO kan redusere Taq polymerasens aktivitet med opp til 50% (Gelfand 1989). DMSO reduserer trolig sekundærstrukturer, og er spesielt nyttig ved G/C rike templatere. Flere sett er tilgjengelig for bruk av DMSO og Betaine i PCR reaksjoner (Frackman et al., 1998). Betaine og DMSO ble tilsatt i PCR reaksjonen før sekvensering.

3.8 Søk etter gen ved hjelp av T7 og genspesifikke forover primere

Det ble prøvd å isolere ønsket innskudd fra biblioteket ved hjelp av T7 vektorspesifikk primer og genspesifikke forover primere direkte på bibliotek. Dette ville gi fragmenter av ulike lengder, der fragmentet ville gå fra begynnelsen av *groEL1* og ut innskuddet. Ut ifra der den genspesifikke primeren bandt seg i den kjente *groESL* sekvensen kunne man

beregne ønsket lengde på PCR produktet. Et bånd på >1500 bp ville være av interesse. Flere reaksjoner ble satt opp der reaksjonsforhold og betingelser var:

Reagenser:	Konsentrasjon:	Volum:
10x Pfu polymerasebuffer	10x	5 µl
10mM dNTPmix	10mM	1 µl
T7 primer		1 µl
For 1		1 µl
Bibliotek		2 µl
Pfu polymerase		1 µl
dH ₂ O til totalvolum		50 µl

Program:	Temperatur:	Tid:	Antall sykluser:
Initiell denaturering	94 °C	5 min	x 1
Denaturering	94 °C	1 min	
Annealing	58 °C	30 min	x 30
Elongering	72 °C	1 min	
Siste elongering	94 °C	5 min	x 1
Hold	4 °C	∞	

PCR produktene ble analysert på agarosegel. Fragmenter med størrelse over 1500bp ble kuttet ut av gelen under UV lys og renset som beskrevet i punkt 3.3.6. Konsentrasjonen av gelrenset DNA kan ofte være noe lav for å kunne sekvenseres. Det ble derfor utført ny PCR på de rensede fragmentene. Å bruke samme sett med primere som ble brukt til å isolere fragmentet kan være problematisk fordi primerene må bindes helt i endene av sekvensen. Forover primer for2 ble derfor benyttet da denne primeren ville binde seg lenger inn enn for1 primeren. Det ble utformet en ny vektorspesifikk primer som bandt seg litt oppstrøms for T7, VektorR. Produktene til disse primersettene ble sendt til sekvensering hos MWG.

3.9 Isolering av *groES1* og *groEL1*

Da operonet *groESL1* var ferdigsekvensert, skulle de to genene isoleres til bruk i ekspresjonsforsøk. Dette ble gjort ved hjelp av PCR og genspesifikke primere utformet til TOPO kloning. Primersettene ble brukt på kromosomalt DNA og med bibliotek som templat. Det kromosomale DNAet som ble benyttet hadde en konsentrasjon på ca. 250 ng/µl. VENT polymerasen som er en korrekturlesende, varmestabil DNA polymerase som lager butte ender på PCR produkter ble benyttet. Det er viktig at PCR produktene som benyttes til TOPO kloning har butte ender.

Reaksjonene ble satt opp:

Reagens:	Konsentrasjon:	Sluttkonsentrasjon:	Volum:
Templat*			1-2 µl
PCR Buffer	10x	1x	10 µl
dNTP mix	10mM	200 µM	1 µl
Primer forover**	10 µM	0.4 µM	2 µl
Primer revers***	10 µM	0.4 µM	2 µl
MgSO ₄	100 mM	4 mM	4 µl
Vent polymerase	2 U/µl	0.02 U/µl	1 µl
dH ₂ O til et totalvolum på			50 µl

*DNA: 250 ng/ µl, volum: 1 µl

Bibliotek: volum: 2 µl

**forover primere til *groES*: ES1forTOPO
groEL: EL1forTOPO
 ***revers primere til *groES*: ES1rew og rew1
groEL: GroEL1R

Fordi de to genene hadde forskjellig lengde, ble 2 ulike programmer brukt for å hente dem opp:

PCR programmene som ble brukt:

Program:	Temperatur:	Tid:	Antall sykluser:
Initiell denaturering	94 °C	5 min	x 1
Denaturering	94 °C	30 sek	
Annealing	62*/64** °C	30 sek	x 25*/30**
Elongering	72 °C	30 sek*/1 min**	
Siste elongering	94 °C	5*/10** min	x 1
Hold	4 °C	∞	

*for *groES*

**for *groEL*

10 µl av hver av PCR prøvene ble analysert på 1 % agarosegel. Forventede bånd for de ulike primersettene var 320 bp for ES1forTOPO+ES1rew, 660 bp for ES1for+rew1 og 1664 bp for EL1forTOPO+GroEL1R.

Båndet av riktig størrelse ble sendt til sekvensering for å undersøke at riktig sekvens var amplifisert. Det var også viktig å se at sekvensene ikke inneholdt mutasjoner, da dette kunne virke inn på ekspresjonsforsøkene.

Det ble også utformet forover primere uten CACC-TOPO hale for å isolere de to genene i operonet, GroES1F og GroEL1F for henholdsvis *groES* og *groEL*. De samme revers primerene ble brukt. Primersettene ble brukt på kromosomalt DNA fra *C. aurantiacus*.

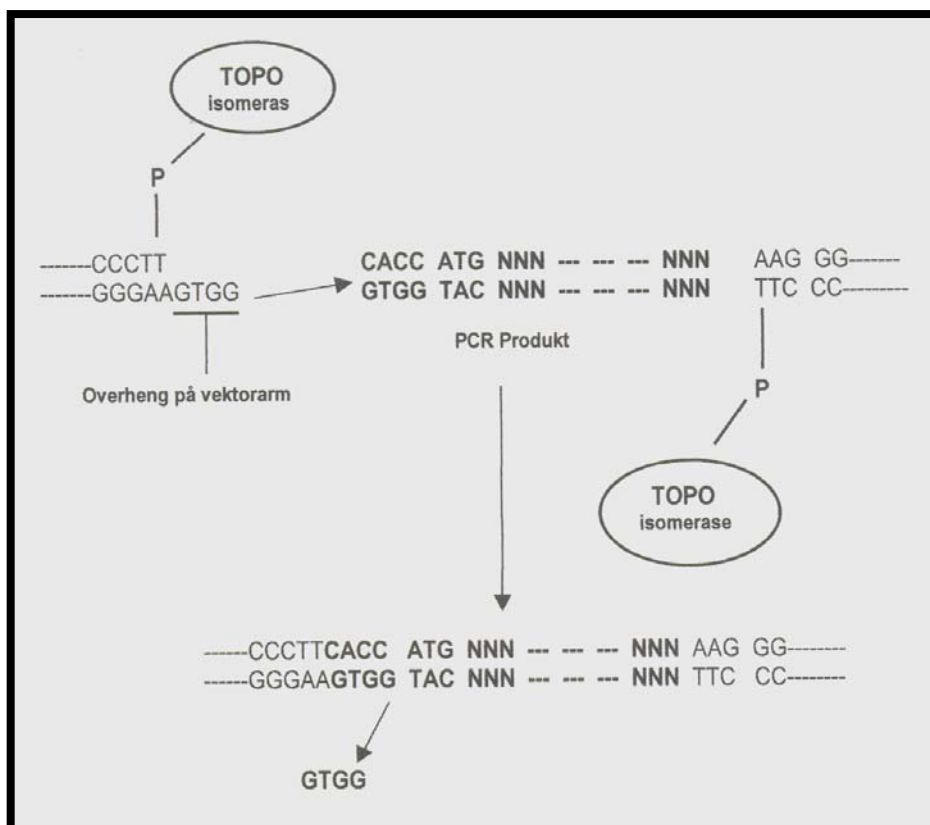
3.10 Transformering og ekspresjon av *groES1* og *groEL1*

3.10.1 Prinsippet for retningsbestemt TOPO-kloning

pET vektorene ble originalt konstruert i 1986 (Studier and Moffatt 1986; Rosenberg et al. 1987, Studier et al. 1990). pET systemet har i dag over 35 typer vektorer og 11 forskjellige vertsstammer av *E. coli*. PET (Plasmid for Expression by T7 polymerase) ekspresjonssystemer bruker elementer fra bakteriofag T7 for kontroll av heterologe gener i *E. coli*. pET TOPO vektoren inneholder promoteren T7, spesifikk for T7 RNA polymerase. Denne promoteren regulerer ekspresjonen av det innsatte genet. Promoteren T7 er modifisert slik at den inneholder en Lac-operator sekvens, der en LacI-repressor kan binde og hindre ekspresjon. T7 RNA polymerase er spesifikk for T7 promotere og gjenkjenner ikke DNA fra andre kilder, da disse sekvensene er sjeldne. T7 RNA polymerase er rundt 5 ganger raskere enn *E. coli* RNA polymerase slik at gener som er under kontroll av T7 promotere kan overuttrykkes. I det T7 baserte ekspresjonssystemet vil T7 RNA polymerase syntetisere mRNA raskere enn *E. coli* RNA polymeraser. Transkripsjon fra T7 promoteren er dermed ikke koblet til translasjon i *E. coli*, noe som resulterer i at mRNA transkriptene ikke blir beskyttet av ribosomene. Transkriptene blir derfor lettere degradert av endogene Rnaser. BL21 Star cellene inneholder en mutasjon som koder for RnaseE (*rne131*) som er en av hovedkildene til mRNA degradering. BL21 stammene gir økt stabilitet av mRNA transkriptene og øker proteinekspresjonen.

Vertsstammen for ekspresjon, BL21 Star™(DE3), bærer i kromosomet en DE3 bakteriofag λ lysogen som har et lac konstrukt med komponentene *lacI* som koder for LacI-repressor og *T7 RNA polymerase* som er under kontroll av lac promoteren. LacI-konstruktet er satt inn i *int* slik at dette er inaktivert. Avbryting av *int* hindrer eksisjon av fagen og dermed lysis. Lac repressoren, lac I, hindrer ekspresjon av både T7 RNA polymerase i vertskromosomet og ekspresjon av det innsatte genet i pET vektoren, ved at repressoren binder seg til operatorsekvensen i lac promoteren og T7 promoteren. Det vil alltid forekomme litt ekspresjon av T7 RNA polymerase fra lac promoteren som igjen vil uttrykke det innsatte genet. Ved å sette en Lac-operator sekvens i T7 promoteren vil all ekspresjon av det innsatte genet forhindres. Selv om ekspresjon av det innsatte genet generelt sett ikke er noe problem, kan det føre til ustabilitet i plasmidet og/eller celledød dersom produktet av det innsatte genet er giftig for *E. coli*. I det opprinnelige lac operonet er det laktose som binder og inaktiverer LacI-repressoren. I pET systemet brukes en kunstig inducer, IPTG, som binder og inaktiverer LacI-repressoren. T7 RNA polymerase vil dermed uttrykkes og feste seg på T7 promoteren slik at det innsatte genet uttrykkes.

Vektoren pET TOPO er linearisert. I hver ende er det bundet topoisomerase. Dette er en topoisomerase I fra *Vaccinia* viruset som binder til spesifikke seter på dobbeltrådet DNA og spalter fosfodiester bindinger etter 5'-CCCTT (Schuman, 1991). Energien fra de spaltede fosfodiesterbindingene brukes til å danne kovalent binding mellom 3'-fosfat i den spaltede tråden og tyrosylenheten i topoisomerase I. Topoisomerasen kan frigjøres og reaksjonen reverseres ved at 5'-hydroksyl-enden på den originale tråden løsner fosfo-tyrosyl-bindingen.



Figur 3.4. Bildet viser ligeringen av PCR produktet inn i pET 102/D TOPO vektoren.

Retningsbestemt TOPO® kloning er en kloningsstrategi som tillater kloning av PCR produkter med et 4 baser overheng i en vektor i 5'-3' retning. Vektoren har et enkelttrådig overheng på 5'-ende, GTGG, og en blunt ende i 3'-ende. Ved å lage PCR produkter med et enkelttrådig 3'-CACC overheng, vil PCR produktet kunne ligere til vektoren (figur 3.4). Det er derfor viktig på forhånd å utforme en forover primer med 5'-CACC begynnelse. Primeren må deretter fortsette rett på ATG, startkodonet til genet man vil klon. Topoisomerasen vil ligere PCR produktet til vektoren. Dersom det er ønskelig å tilføre en hale til den C-terminale enden, V5 epitop og histidinhale, kan reversprimere til dette formålet utformes. I denne oppgaven ble foroverprimerene ES1forTOPO og EL1forTOPO utformet med 5'-CACC overheng. Revers primerene ble utformet slik at geneses native stoppkodon ville terminere transkripsjonen.

PCR produktene som var amplifisert med de genspesifikke TOPO-primerene og VENT polymerasen ble brukt i TOPO kloningen. Disse produktene hadde 5'-CACC enkelttrådig overheng og butte ender, og kunne derfor brukes direkte til kloning i pET101/D-TOPO og pET102/D-TOPO vektorene. Denne vektoren kloner optimalt når det molare forholdet mellom innskudd og vektor er mellom 2:1 og 5:1. Dette var anbefalt av produsenten, og ble regnet ut etter formelen:

$$\frac{\text{ng vektor} \times \text{størrelse innskudd (bp)} \times \text{molar ratio(innskudd/vektor)}}{\text{Størrelse vektor (bp)}} = \text{ng innskudd}$$

Innskuddet ble satt inn i vektoren som beskrevet i TOPO manualen. Da *groES* var på 320 bp og *groEL* på 1664 bp, ble tiden på liggeringen satt til henholdsvis 10 minutter og 30 minutter, etter produsentens anbefaling. Figur 3.5 viser hvordan innskuddet liggeres inn i TOPO vektoren ved hjelp av topoisomerase.

3.10.2 Transformerings av *groES1* og *groEL1*

Transformeringsen ble utført etter protokollen til One Shot TOP10 kompetente celler, som var det settet som ble benyttet. I dette settet overføres cellene, etter transformasjon, til et rikt medium, S.O.C., der de stabiliserer seg. Dette mediumet er ikke selektivt, så det er viktig at cellene ikke oppbevares i dette mediumet så lenge at de begynner å vokse (ca. 1 time). Deretter strykes 100-200 µl ut på LB skåler tilsatt ampicillin (100 µg/ml) og inkuberes over natt ved 37 °C i varmeskap.

Da det kun er seleksjonsmarkør som forteller om cellene har tatt opp plasmid, men ikke om plasmidet har innskudd eller ikke, må koloniene som vokser opp på de selektive LB skålene analyseres. Dette ble gjort ved å kjøre PCR direkte på cellene.

Reaksjonen ble satt opp:

Reagens	Konsentrasjon:	Sluttkonsentrasjon:	Volum:
DynazymeEXT buffer	10x	1x	5 µl
dNTPmix	10mM	200 µM	1 µl
templat	litt av en koloni, plukkes med tannpirker		
T7	10 µM	0.4 µM	2 µl
T7revers	10 µM	0.4 µM	2 µl
DynazymeEXT polymerase			1 µl
dH ₂ O til et totalvolum på			50 µl

Primersettet T7 og T7revers er plasmidspesifikke primere som binder seg henholdsvis et lite stykke oppstrøms og nedstrøms for kloningssetet. PCR på plasmider uten innskudd vil med disse primerene gi produkt 632 bp for pET 102/D TOPO vektor. Ut ifra båndenes størrelse på gelen kan man skille plasmider med innskudd fra plasmider uten innskudd. Denne størrelsen må også legges til når man beregner størrelsen til innskuddet. Koloniene ble plukket med en steril tannpirker og skyllet i PCR blandingen. Det er viktig å merke av på skålen hvilken koloni man plukker slik at man kan gå tilbake og plukke flere celler fra denne kolonien dersom den viser plasmid med innskudd.

PCR programmet som ble brukt:

Program:	Temperatur:	Tid:	Antall sykluser:
Initiell denaturering	94 °C	10 min	x 1
Denaturering	94 °C	1 min	
Annealing	57 °C	30 sek	x 25
Elongering	72 °C	30 sek*/1 minutt**	
Siste elongering	94 °C	5 min	x 1
Hold	4 °C	∞	

*for *groES*

**for *groEL*

Prøvene ble analysert på agarosegel ved elektroforese. Bånd på størrelsen 952 bp for *groES* og 2296 bp for *groEL* tydet på plasmider med innskudd, og kolonien til disse plasmidene ble dyrket opp i flytende LB medium tilsatt ampicillin (100 µg/ml) og inkubert over natt ved 37 °C i risteinkubator. Plasmidet pET har et gen for ampicillinresistens som koder for β-laktamase. β-laktamase skiller ut i mediet der det hydrolyserer og inaktiverer ampicillin. Ampicillin fjernes derfor fra mediet etter hvert slik at ikke-selektive forhold oppstår, og celler uten vektor vil derfor kunne vokse opp. Dette kan resultere i at det ønskede genet ikke blir uttrykt og at celler uten plasmid selekteres.

For å kunne transformere BL21 Star™(DE3) *E. coli* cellene med pET TOPO konstruktet, må man ha rent plasmid-DNA. Plasmidene fra overnatt-kulturene ble rensset ved hjelp av Wizard® Plus SV Minipreps DNA Purification System Kit (Promega) som beskrevet i tilhørende protokoll. Metoden er effektiv for å isolere plasmid-DNA og går ut på å åpne cellene med NaOH og SDS. Kromosomalt DNA og proteiner danner uløselige aggregater, mens plasmid-DNA er løst i supernatanten. Plasmid-DNAet i supernatanten spinnnes ned i en spinkolonne mens kromosomalt DNA og proteinaggregater blir igjen i silika spinnfilteret. De rensede plasmidene analyseres på gel med kvantitative og kvalitative standarder. Ut ifra gelen kan konsentrasjonen til plasmidene bestemmes. For å analysere de isolerte plasmidene for innskudd ble det også utført en PCR med vektorspesifikke primere. Programmet som ble brukt var det samme som ble brukt da koloniene ble testet for innskudd. PCR prøvene ble analysert på agarosegel.

3.10.3 Ekspresjon av *groES1* og *groEL1*

Kjemisk kompetente BL21 Star™(DE3) celler ble transformert med pET/D-TOPO vektorer med innskudd som beskrevet i protokollen. Før cellene ble tilsatt IPTG ble de inkubert 2 timer ved 37 °C i LB-medium tilsatt ampicillin. Konsentrasjonen av IPTG var 0,5 mM. En negativ kontroll, ikke tilsatt IPTG, ble alltid inkubert parallelt. For å forsikre om at reaksjonene gikk som de skulle, ble en positiv kontroll alltid benyttet. Kontrollproteinet var LacZ fra kontrollvektoren pET 101/D/lacZ-TOPO og pET 102/D/lacZ-TOPO. Toksiske proteiner kan være problematiske å få uttrykt, og ekspresjonen må ofte reguleres. Ekspresjon av genene i pET/D/TOPO vektorene kan reguleres ved å senke veksthastigheten til BL21(DE3) cellene. Dette kan f. eks. gjøres ved å senke inkubasjonstemperaturen fra 37 °C til romtemperatur, eller ved å tilsette 1 % glukose i mediet.

Etter induksjon ble det tatt prøver hver time i til sammen 5 timer. Prøver på 500 µl ble tatt ut, spunnet ned (10000rpm/10min), supernatanten ble kastet og cellene oppbevart på – 20 °C.

Da det av resultatene kom frem at *groEL1* ikke lot seg uttrykke med verken pET 101/D/lacZ-TOPO eller pET 102/D/lacZ-TOPO, ble nukleotidsekvensen undersøkt i programmet Rare Codon Calculator for kodoner som sjeldent brukes i *E. coli*. Da dette viste seg at sekvensen inneholdt 14 sjeldne kodoner (punkt 4.7) ble en vertscelle, BL21-CodonPluss (DE3) RP forsøkt. BL21-CodonPlus (DE3) RP vertscellen inneholder ekstra kopier av *argU* og *proL*, genene som koder for tRNA som gjenkjenner kodonene for arginin; AGA og AGG, og kodonet for prolin; CCC. BL21-CodonPlus (DE3) RP vertscellen ble transformert med pET 101/ D TOPO vektoren og dyrket på LB skåler som inneholdt ampicillin og cloramfenikol. Genene som koder for ekstra tRNA ligger på et plasmid som også bærer kloramfenikol-resistens. Cellene ble indusert på samme måte som BL21 Star™(DE3) vertscellene.

I tillegg til positiv kontroll ble negative kontroller alltid brukt. Dette var vertsceller med plasmid uten innskudd og vertsceller uten plasmid.

3.11 SDS-PAGE

SDS-PAGE er en vanlig metode for å separere proteiner i størrelsesområdet 15-200 kDa. Metoden baserer seg på separasjon av proteiner avhengig av proteinets størrelse og brukes til å kvantifisere, sammenligne og karakterisere proteiner. SDS-PAGE utføres med en todelt gel i et diskontinuerlig buffersystem. Gelene lages av polyakrylamid som er en kryss-linket polymer av akrylamid. Lengden av polymerene avhenger av konsentrasjonen av akrylamid, som typisk ligger mellom 3,5-20 %. Slike geler er mer tidkrevende å lage enn agarosegeler, da de er todelte og inneholder flere reagenser. I tillegg inhiberes polymeriserings prosessen av oksygen og gelene må derfor lages mellom to glassplater eller i sylindre. Polyacrylamide kan brukes til å separere DNA fragmenter mindre enn 500 bp, men brukes først og fremst til proteinseparering. Den øverste delen av gelen har lav akrylamidkonsentrasjon (5 %) og er en konsentrerende gel med lav ionisk styrke og pH omtrent nøytral. I denne gelen konsentreres prøvene før den når frem til den andre delen av gelen. Den nederste delen av gelen har høy akrylamidkonsentrasjon (7-15 %), høyere ionisk styrke enn den øverste gelen og en pH på rundt 8. Gelene som ble benyttet i denne oppgaven ble laget som beskrevet i Bollag *et al* (1996).

Prøvene som ble analysert på SDS-PAGE var celler fra induksjonsforsøket med BL21 Star™(DE3). For å ødelegge celleveggen ble cellene sonikert. Sonikering forårsaker mekanisk ødeleggelse av celleveggen og fører til at cellematerialet blir frigjort i løsningen. Cellene ble løst i 20 mM kaliumfosfatbuffer, pH = 7,5, 1 ml buffer per gram cellemateriale. Prøvene ble sonikert på is, 2 x 5 sekunder. Rørene ble deretter sentrifugert (13000rpm/30min/4°C). Supernatanten som inneholdt proteiner ble satt på is til videre analyser, eller frosset ned ved - 20 °C. Mengde protein i prøven må bestemmes før prøvene kan appliseres på gelen. Dette ble gjort etter Bradford-metoden. Bio-Rad Protein Microassay fra Bio-Rad og den tilhørende protokollen ble fulgt. Absorpsjonsmaksimumet for i en sur Coomassie Brilliant Blue G-250 løsning skifter fra 465 nm til 595 nm når denne løsningen bindes til protein. Løsningen binder seg hovedsakelig til basiske og aromatiske aminosyrer. Fargen blå vil vise tilstedeværelse av proteiner, rød tyder på at proteiner ikke er tilstede. Prøvene kan eventuelt fortynnes med kaliumfosfatbuffer, pH 7,5. Metoden er følsom for 1-20 µg/ml proteinløsninger. BSA, Bovint Serum Albumin, ble benyttet som standardprotein og absorbansen målt spektrofotometrisk ved 595 nm som blindprøve. En BSA standardkurve konstruert av Ingun Benestad (Hovedoppgave UIO, 2005) ble benyttet til utregning av proteinmengdene.

Prøver som skal kjøres med SDS-PAGE må først kokes i en prøvebuffer som inneholder glyserol, β-merkaptøetanol og SDS. Glyserolen i prøvebufferen gjør at prøvene får en viskositet som egner seg godt når prøvene skal pipetteres i brønnene i polyakrylamidgelen. SDS er en anionisk detergent og binder sterkt til hydrofobe deler av proteinet og denaturer dem. β-merkaptøetanol reduserer disulfidbroene som holder den tertiære proteinstrukturen sammen. Hvert protein i blandingen blir åpnet til en stavformet struktur med en serie negativt ladete SDS molekyler bundet langs polypeptidkjeden og på den måten fullstendig denaturert. Ladningen blir proporsjonal med polypeptidets

lengde og molekylet vil vandre i polyakrylamidgelen med en relativ mobilitet avhengig av molekylets lengde. Molekylene vil skilles i gelen etter størrelse, da små molekyler vil bevege seg lettere gjennom porene i polyakrylamidgelen. Prøvebufferen inneholder også fargestoffet bromfenolblått som synliggjør vandringsen til proteinene i gelen.

Prøvene med protein kokes med bufferen på vannbad i 5 minutter før de avkjøles og pipetteres i brønnene på gelen. Hver brønn ble tilsatt omtrent 20-100 µg protein, avhengig av forsøket. Elektroforesen foregår ved 120 V i ca. 1 time. Deretter tas gelen ut av glassplatene, farges i ca. 20 ml Coomassie Blue fargeløsning i 5-10 minutter på ristmaskin ved romtemperatur. Fargeløsningen skal deretter fjernes og gelen skylles i vann. Deretter skal gelen stå med ca. 50 ml Coomassie avfargeløsning. Tiden avfargingen tar avhenger av båndene som skal synliggjøres, 1 time for svake bånd, overnatt for meget sterke. Proteinene på gelen sammenlignes med størrelsesmarkøren som er applisert i en av brønnene på gelen. Gelen kan fotograferes under hvitt lys.

Størrelsen på gelene som ble brukt i denne oppgaven var 10 cm. x 7,5 cm. x 0.75 mm. Precision Plus Protein™ Standards Unstained fra Bio-Rad ble brukt som størrelsesmarkør.

3.12 Sekvensering av *groEL1*

Da det viste seg å være vanskelig å få *groEL1* sekvensert, ble sekvensen *in silico* analysert i programmet mfold. Det viste seg at sekvensen inneholdt tredimensjonale strukturer som kunne tenkes å skape problemer for sekvenseringen. Det ble derfor forsøkt å redusere dannelse av sekundærstrukturer ved å bruke et eget sekvenseringsprogram. Programmet ble satt opp i samarbeid med avdelingsingeniør Bård Mathisen.

Sekvensering av nukleotidsekvenser foregår ved 44 °C på enkelttrådig DNA. Dobbelttrådig templat gjøres enkelttrådig i PCR ved bruk av en primer. Denne PCR reaksjonen foregår som en vanlig PCR, med 3 trinn. Ved å hoppe over annealingstrinnet, og utføre en to-trinns PCR istedenfor kan man holde temperaturene høyere gjennom hele reaksjonen. På denne måten kan man minke dannelsen av sekundærstrukturer.

Det ble satt opp 3 parallelle programmer:

1.	Denaturering	30 sek	95 °C
	Elongering	60 sek	60 °C
2.	Denaturering	30 sek	95 °C
	Elongering	120 sek	60 °C
3.	Denaturering	30 sek	95 °C
	Elongering	240 sek	60 °C

Elongeringen og denatureringen ble repetert med 30 sykluser. For alle 3 programmene ble det brukt en "hot-start" på 5 minutter ved 95 °C. Det ble ikke brukt et siste elongeringstrinn.

Reagenser:	Volum:
Templat PCR produkt	5 µl
Exozap-IT	2 µl
Dynamic ET-dye	8 µl
Primer 5 µM	1 µl
dH ₂ O	4 µl
Totalvolum	20 µl

Primerene GroEL1F og GroEL1R ble brukt i hvert program, 2 paralleller, i alt 12 reaksjoner. Templatet var PCR produkt av primersettet GroEL1F/GroEL1R på 1664 bp. Templat og Exozap-IT ble kjørt på et program, enzym 30, før resten av reagensene ble tilsatt. Exozap-IT renses templatet for PCR rester, inaktiverer Mg²⁺ og fjerner salter. Programmet enzym 30 ble kjørt i PCR maskin og bestod av 15 minutter på 95 °C og 30 minutter på 70 °C. Deretter ble primer, dH₂O og Dynamic ET-dye tilsatt og de tre programmene kjørt. Dynamic ET-dye er en reaksjonsmiks som inneholder polymerase, dNTP og buffer.

Av PCR-produktene ble 5 µl analysert på agarosegel. Ethidiumbromid vil kunne interkalere med enkelttrådig DNA i områder med sekundærstrukturer. Dette vil være synlig på en agarosegel tilsatt ethidiumbromid, belyst med UV-lys. 5 µl av PCR produktet ble også inkubert på vannbad, 44 °C i 4-6 timer. Sekvenseringsreaksjonen foregår ved denne temperaturen og varer i over 4 timer, avhengig av strømstyrken sekvenseringen kjøres på. PCR produktene ble deretter analysert på agarosegel. Dette ble gjort for å se om de sekundære strukturene til det enkelttrådig DNAet oppstod i løpet av sekvenseringen ved 44 °C eller under PCR reaksjonen. Sekvenseringen ble utført ved 44 °C og på 50 °C for å prøve å forhindre dannelse av sekundærstrukturer.

Som kontroll ble det laget enkelttrådig DNA etter de vanlige betingelsene for sekvensering av PCR produkter. Reagensene var de samme, men programmet som ble brukt var et tre-trinns PCR program:

Denaturering	95 °C	20 sekunder	x 30 sykluser
Annealing	50 °C	15 sekunder	
Elongering	60 °C	60 sekunder	

Produktene ble behandlet på samme måte som produktene fra de andre oppsettene.

3.13 Databehandling av sekvenseringsresultatene

Sekvenseringsresultatene fra MWG og GATC ble behandlet i sekvensbehandlingsprogrammet BioEdit. Sekvensene ble brukt i søk i databaser som NCBI og EMBL-EBI. Ulike typer søk ble gjort avhengig av hvilken informasjon som var av interesse. Første søk på en mottatt sekvens var å bruke den i BLAST der den ble analysert mot genomet til *C. aurantiacus*. Dette for å se om den fikk treff på genet man var ute etter. Var det riktig sekvens ble den sekvenssammenstilt med målsekvensen i programmet Multalin og BLAST, align two sequences. Sekvensen ble deretter brukt til å søke i prokaryote genomer generelt. Det siste av interesse for å finne informasjon om andre chaperoniner og likhetene deres med *groESL1* fra *C. aurantiacus*.

Multalin er et program for sekvenssammenstilling. I Multalin ble de sekvenserte sekvensene sammenstilt med den allerede kjente sekvensen til *groESL1* og *groESL2* for å confirmere at riktig operon var funnet og for å se om den eventuelle fortsettelsen av dette operonet var kommet med.

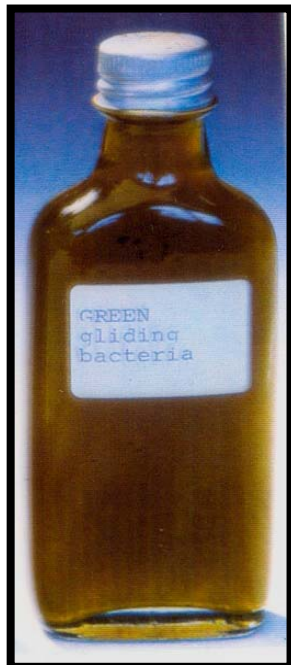
Resultatene fra MegaBACE maskinen ble analysert på samme måte som sekvensene mottatt fra GATC og MWG. Multalin ble brukt i noe større grad da disse sekvensene var isolert ved hjelp av de genspesifikke primerene og ble sekvensert som en kontroll for at riktig klon ble undersøkt videre under søk i biblioteket.

En rekke ulike programmer ble i tillegg brukt til å behandle sekvensene og til å innhente informasjon om dem. GeneDoc ble brukt til fler-sekvenssammenstillinger av både nukleotidsekvenser og aminosyresekvenser, blant annet i forbindelse med søk etter en konsensusprimer. Nukleotidsekvenser ble analysert for tredimensjonale strukturer i mfold. Restriksjonssteder ble funnet med programmene WebCutter og NebCutter. Ved hjelp av programmene PRETTYSEQ og PeptideMass ble aminosyresekvenser utledet og molekylvekter beregnet. Tredimensjonale strukturer i nukleotidesequensene ble funnet ved hjelp av programmet mfold.

Adressen til programmene er å finne i oppgavens del om materialer (2.8). Bruken av programmene og deres verdi for oppgaven er nærmere beskrevet i oppgavens resultatdel.

4 RESULTATER

4.1 Dyrking og høsting av *Chloroflexus aurantiacus*.



C. aurantiacus ble dyrket i rikt medium i 60 ml flasker, som vist på figur 4.1. Flaskene sto i varmeskap ved 55 °C, 20-30 cm fra en lyskilde på 40 W. Cellerekken ble holdt i livet ved at man hver 3de uke overførte celler til friskt medium.

C. aurantiacus er en lett organisme å holde i kulturer dersom forholdene er lagt til rette. Mediet den vokser i er verken tidkrevende eller vanskelig å lage. Ved gode vekstbetingelser kan organismen sees som tette skyer i flasken. Etter ca. 4 dager har cellene som regel oppnådd en tetthet med $OD_{600} = 1$, og det er da de egner seg best til ekstraksjoner. Etter dette vil døde celler falle til bunnen slik at flasken inneholder degradert cellemateriale. Dette cellematerialet egner seg dårlig til DNA isolering da de vil gi mye degradert DNA i prøvene.

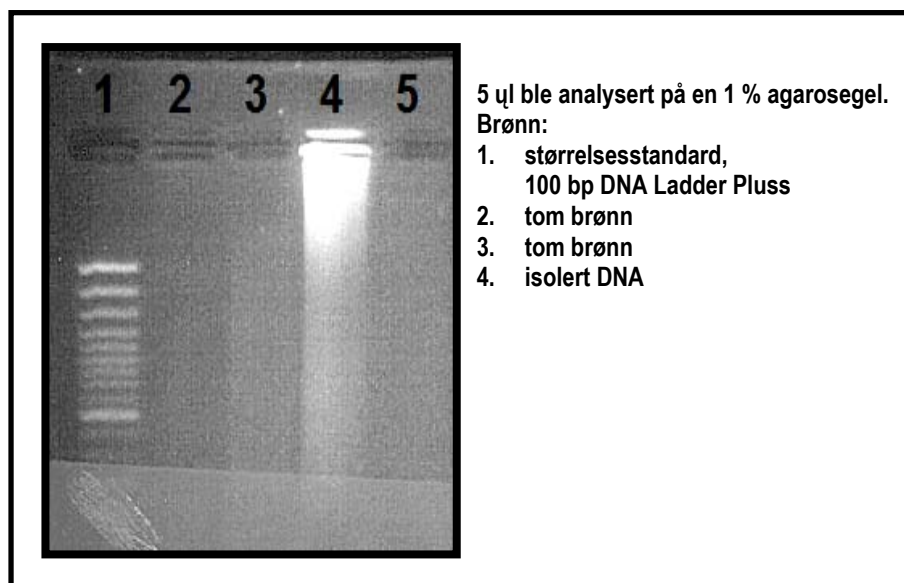
Figur4.1 *C. aurantiacus*

4.1.1 Isolering av kromosomalt DNA, måling av DNA konsentrasjon og undersøkelse av kvaliteten til isolert DNA.

Kromosomalt DNA fra *C. aurantiacus* ble isolert etter fenol/kloroform metoden som beskrevet i delen om metoder. Renheten og kvaliteten på det isolerte DNAet ble bestemt spektrofotometrisk og ved agarose gelelektroforese. Rent DNA vil gi en absorbans ratio av A_{260}/A_{280} på verdier rundt 1,8. DNAet ble analysert ved agarose gelelektroforese. Det isolerte DNAet var noe degradert. På bildet av gelen, figur 4.2, kan degradert DNA kan sees som smørje i brønn 4 med DNA. Smørjen kan skyldes cellenes egne DNaser som vil degradere det kromosomale DNA når celleveggen ødelegges, eller mekanisk kutting under isoleringen. Det isolerte DNAet ble likevel brukt i forsøkene, fordi graden av degraderingen ikke var stor.

Spektrofotometriske målinger og beregninger:

Utgangsmateriale:	1,2 g celler, våtvekt
A_{260}/A_{280} :	$0,358 / 0,208 = 1,721$
Konsentrasjon $\mu\text{l/ml}$:	$50 \times 0,358 \times 14 = 251$
Utbytte:	0,25 mg DNA



Figur 4.2. Isolert kromosomalt DNA fra *C. aurantiacus*.

DNAet som ble isolert ble brukt i PCR, til å undersøke primerenes spesifisitet og til å isolere de to genene *groES1* og *groEL1*.

4.2 Genbibliotek

4.2.1 Test av genbiblioteket

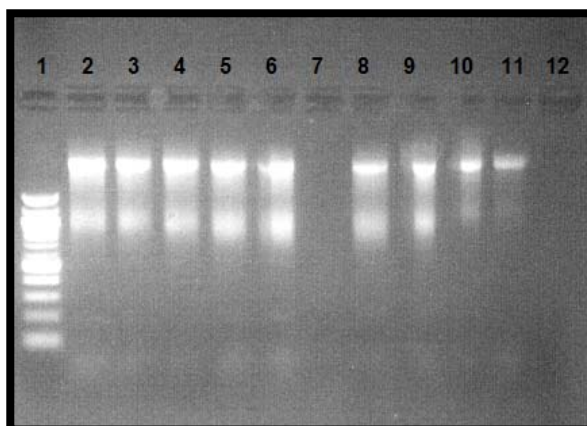
Et genbibliotek i fag λ som ikke hadde vært i bruk siden 2001, ble undersøkt med tanke på om det fortsatt inneholdt innskudd og om fagen fortsatt hadde evnen til å infisere vertscellene.

Ved å utføre blå hvit bakgrunnstest var det mulig å undersøke om store deler av biblioteket eventuelt var uten innskudd. Skålen med celler og fag, dyrket på LB medium tilsatt X-gal og IPTG, ble inkubert ved 37 °C over natt og analysert. Det var en del blå plakk i biblioteket, noe som tydet på at det var fag uten innskudd tilstede. Årsaken til at biblioteket inneholdt fag uten innskudd kan være at biblioteket var amplifisert med fag uten innskudd og/eller at innskuddene var gått tapt fra fagen. Det var allikevel ikke så stor andel av blå plakk at det ville ha betydning for videre arbeid, som vist ved bruk av genspesifikke primere på bibliotek som templat. Blå hvit test ble også utført på positivt fageluut, resultatet viste færre blå plakk.

En test ble også utført ved hjelp av PCR, med biblioteket som templat og primersettet T7/T3, som er de genspesifikke primerene til λ ZAP bakteriofagen. PCR ble utført i Techne gradient maskin, med temperaturgradienten 45-65 °C (figur 2.3).

Som forventet viste produktene fra PCR reaksjonen smørje fordi genomet til *C. aurantiacus* var delvis kuttet med restriksjonsenzymet *Sau3AI*. Delvis kutting av et genom fører til tilfeldige fragmenter av ulik størrelse. Størrelsen på fragmentene som var satt inn i vektoren var på 2-5 kb (Vanberg, hovedoppgave UIO, 2001). Bildet som er vist i figur

4.3 viser en del mindre fragmenter i tillegg, samt svake bånd på ca. 300 bp. Dette er størrelsen på de produktene man vil få med det gitte primersettet i en PCR reaksjon dersom vektoren er uten innskudd. De nederste båndene på gelen er primer-dimer.



Figur 4.3. Gradient PCR på bibliotek med primersettet T7/T3 på bibliotek, 10 μ l av hvert PCR produkt ble analysert på en 1 % agarosegel. Brønn 1: 100 bp DNA Ladder, brønn 2-11: PCR produktene, brønn 12: blindprøve, dvs PCR produkt uten templat DNA. Brønn 7 viser ingen bånd, PCR reaksjonen har her ikke fungert.

Genomstørrelsen til *C. aurantiacus* var antatt å være på omtrent 4×10^6 bp (samme som hos *E. coli*). Dersom man regner at gjennomsnittet på de ligerte fragmentene i biblioteket var omtrent 4 kb, vil i prinsippet 1000 fag representere hele genomet til *C. aurantiacus*. Fordi fag vektoren inneholder tilfeldige fragmenter av det genomiske DNAet, vil noen av disse fagene inneholde de samme fragmentene, eller deler av samme sekvens og noen vil ikke inneholde noe i det hele tatt. For å være sikker på at genomet er representert minst en gang er det nødvendig med omtrent 10 ganger så mange fag i praksis som i prinsippet. Før biblioteket ble amplifisert var totalvolumet 500 μ l og titeret $3,0 \times 10^4$ fag (Vanberg, hovedoppgave UIO, 2001). Antall fag var dermed totalt ca. $1,5 \times 10^6$. Sannsynligheten for at et bestemt DNA fragment skal finnes i et tilfeldig bibliotek med N uavhengige rekombinanter kan beregnes med formelen (Clarke og Carbon, 1976):

$$N = \frac{\ln(1-P)}{\ln(1-1/n)}$$

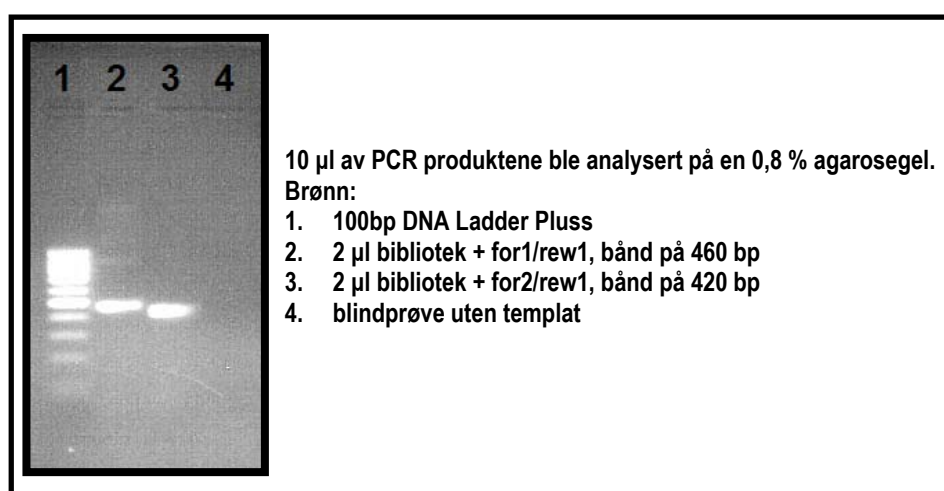
Der N er antall rekombinanter, P er sannsynligheten for at et bestemt fragment er tilstede og n er størrelsen av genomet i forhold til størrelsen på fragmentene som er brukt.

Dersom N er lik $1,5 \times 10^6$ og n lik $4,0 \times 10^6$ bp/4000 bp, som i vårt tilfelle, vil P bli lik 1. Ut i fra dette er det teoretisk sett 100 % sannsynlighet for at det ønskede fragmentet er tilstede i biblioteket. Biblioteket vil dermed representere hele bakteriens DNA dersom genomet har den antatte størrelsen på rundt 4,0 kb.

4.2.2 Optimalisering av primere til søk i biblioteket etter positiv fagklon

Primersettene som skulle brukes i søk etter positive plakk i bibliotek med metodene til King og Yu og Bloem, ble utformet av stipendiat Marc Anglès d'Auriac. *GroESL1* og 2 ble sekvenssammenstilt og primere utformet i områder der sekvensene var ulike. Det ble utformet 2 sett primere i første omgang, for1/rew1 og for2/rew1, som ville isolere fragmenter på henholdsvis 460 og 420 bp. Før de ble brukt i søk ble de optimalisert på biblioteket. De ble også prøvd på kromosomalt DNA.

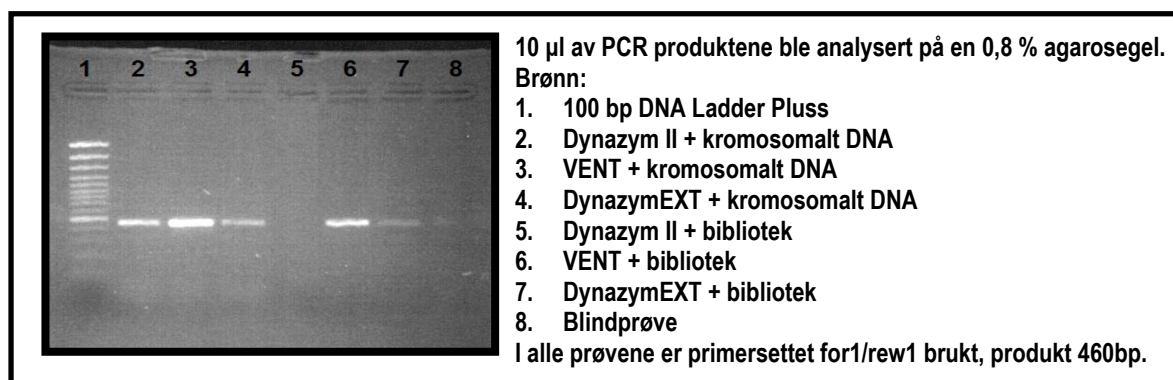
En annealingstemperatur på 57 °C viste seg å være noe lav og ble etter optimalisering med Techne gradient PCR maskin økt til 62 °C. Ulike mengder templat ble prøvd ut. Mengde bibliotek, på 1 µl, viste seg å gi indikatorbåndet av og til, mens 2 µl ga produkt av forventet størrelse. Bildet på figur 4.4 viser PCR produktene etter optimalisering.



Figur 4.4. Etter optimalisering av reaksjonen.

Primersettet for1/rew1 viste seg å alltid gi bånd på 460 bp når det ble brukt i PCR under optimaliserte betingelser på 2 µl bibliotek som templat. Dette ble derfor brukt som positiv kontroll i hver PCR reaksjon under søk på biblioteket. Primersettet for1/rew1 ble brukt på alle skålene og sektorene, også på de positive plakkene. En negativ kontroll uten templat DNA ble alltid brukt.

På figur 4.5 vises et bilde av PCR produkter der samme reaksjonsbetingelser er brukt, med samme PCR program, men med ulike polymeraser.



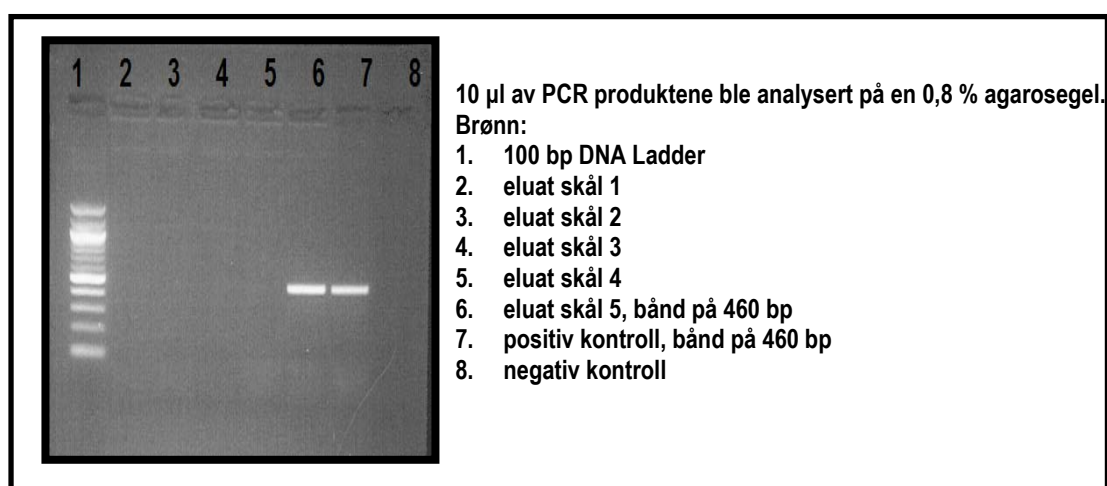
Figur 4.5. Test av ulike polymeraser på kromosomalt DNA og på genbiblioteket.

Reaksjonen viste at VENT ga best resultat og denne polymerasen ble derfor brukt i arbeidet videre med genbiblioteket. I tillegg til mengde templat og ulike polymeraser ble også annealingstemperaturen optimalisert ved hjelp av Techne gradient PCR maskin. Denne viste at reaksjonen ga best og mest spesifikke resultater på 62 °C.

For å forsikre om at det var riktig fragment som ble oppformert i PCR reaksjonen ved bruk av for1/rew1 primersettet, ble produktet sendt til sekvensering. Fordi nettopp dette primersettet skulle brukes til å identifisere fag med et spesifikt innskudd, var det svært viktig at primerene hentet opp det nøyaktige fragmentet. Det var også viktig å finne ut at det var *groESL1* primersettet festet seg til, og ikke *groESL2*. Produktet var på 460 bp og ble sekvensert. Resultatet ble oppgitt i form av et kromatogram. Informasjonen ble analysert i Bioedit og oversatt til fasta format. Sekvensen ble brukt til søk i NCBI's BLAST program der den ble sammenstilt mot genomet til *C. aurantiacus*, og i Multalin, der den ble sekvenssammenstilt med *groESL1*. Resultatet viste seg å være korrekt. Kromatogrammet er vedlagt Appendiks.

4.2.3 Søk etter positiv fagklon i genbiblioteket

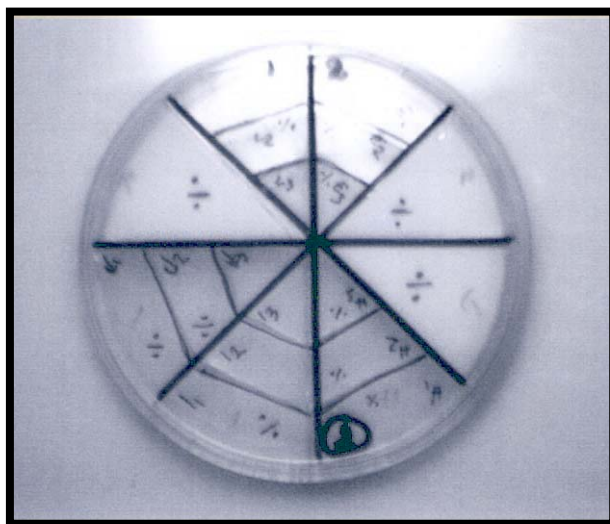
Metoden til Yu og Bloem, 1996, ble noe modifisert, og brukt i kombinasjon med metoden til King, 1997. LB skåler med XL1 Blue MRF⁻ celler og fag fra biblioteket ble dyrket opp som beskrevet i delen om metoder. Disse skålene ble først analysert etter metoden King. På disse skålene var det omkring 10 000 plakk. Positive primære subbibliotek fra metoden etter King ble brukt til å så ut et sekundæresøk på nye skåler med en plakkmengde på omkring 200 plakk per skål. Ved utregning av titeret til biblioteket og ved fortynningsrekker ble det klart at ca. 200 plakk ville gi godt isolerte plakk som ville være lette å plukke dersom noen av dem skulle vise seg å være positive. Dette var viktig da enkle plakk skulle isoleres, og fordi det var viktig at to plakk ikke gikk over i hverandre og forstyrret analysene. Skåler med positive plakk ble analysert videre etter metoden til Yu og Bloem. Mulige plakk ble snevret inn etter hvert som sektorene ble merket positivt. Til slutt var det tilbake omkring 20 potensielt positive plakk som kunne isoleres og analyseres.



Figur 4.6. Primærsøk etter metoden King.

Resultatene fra PCR på de primære subbibliotekene viste hvilke skåler som inneholdt positive kloner. Bildet på figur 4.6 viser resultatet fra 5 av skålene som ble testet. I brønn 6 ser man indikatorbåndet på 460 bp, som viser at denne skålen inneholder plakk med innskudd av interesse. Eluatet fra denne skålen ble brukt til å lage nye skåler, med isolerte plakk som ble brukt videre med Yu og Bloems metode. Brønn 7 viser at reaksjonen har gått som den skal ved et klart bånd i positiv kontroll og brønn 8 at det ikke har vært noen forurensninger i prøvene.

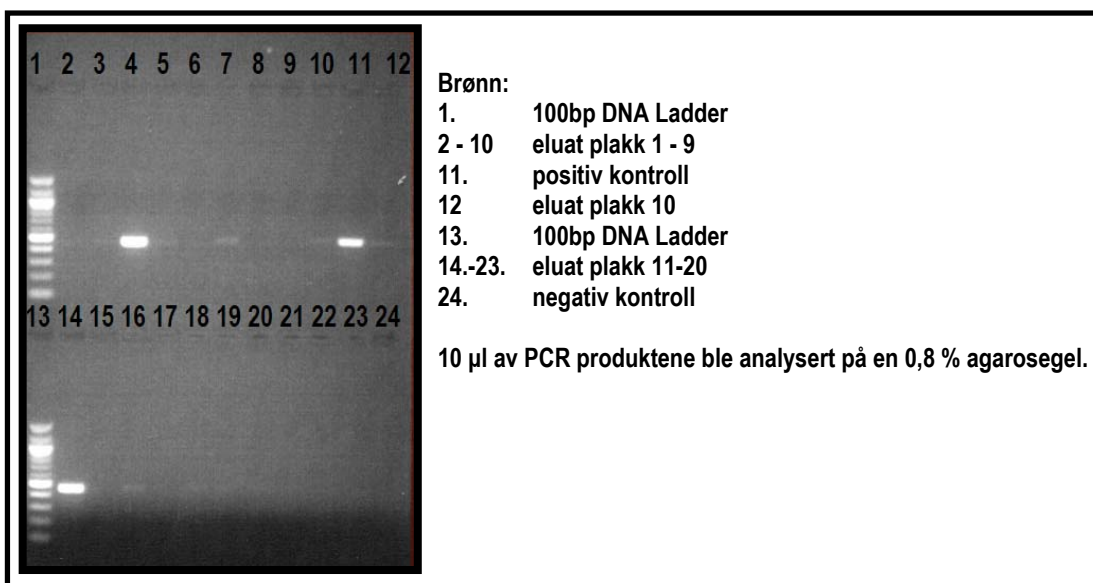
Ifølge metoden til Yu og Bloem skulle en membran legges over skålene. Til membranen festet det seg celler og plakk som ble skylt av i SM-buffer. SM-bufferen, eluatet, ble brukt som templat i PCR. For at ikke toppagaren skulle følge med når membranen ble løftet opp, ble det brukt agarose i toppagaren istedenfor agar. For å undersøke en positiv skål, ble en ny membran lagt over. Membranen ble delt inn i sektorer ved hjelp av en steril skalpell, og hver membranbit skyllet separat i SM-buffer. Eluatet ble brukt som templat i PCR. På denne måten ble skålen snevret inn i mindre sektorer etter hvert som positive sektorer ble identifisert. Dette er vist på fotografiet på figur 4.7.



Figur 4.7. Fotografi av en agarskål med sektorsøk etter positiv plakk etter metoden til Yu og Bloem. Bildet viser inndeling av sektorene og er tatt i gelbildeapparat, med hvitt underlag og hvitt lys. Sektorene som ikke ga positivt resultat er merket -. Sektorer som ikke er merket, undersøkes videre.

Ifølge Yu og Bloem skulle membranen som tilhørte den positive underseksjonen hybridiseres med en probe for å plukke ut en enkelt positiv plakk. Sektorene på skålene ble snevret inn til det var omkring 20 plakk per sektor. Disse ble istedenfor plukket med en steril glass pipette og eluert i 50 µl SM buffer. Eluatet ble brukt i PCR reaksjonen som beskrevet i delen om metoder. Det ble også forsøkt å plukke plakken ved hjelp av en steril tannpirker som deretter ble skylt i PCR-blanding, men denne metoden ga ikke tilfredsstillende resultater.

Bildet på figur 4.8 viser PCR på fageluat fra 20 ulike plakk. Brønn 4 og 14 viser klare positive resultater. Brønn 5, 7, 10, 12, 16 og 18 viser veldig svake bånd. Eluatet til de positive plakkene ble oppbevart på -20 °C. Når det viste seg at plakken var positivt kunne man gå tilbake til denne plakken som var løst i SM buffer. Eluatet ble amplifisert opp etter samme teknikk som brukt for å amplifisere biblioteket, for at det ikke skulle gå tapt og for at det skulle finnes i rikelige mengder. Eluatet fra den originale plakken ble brukt videre i reaksjoner for å hente opp innskuddet med de vektorspesifikke primerene T7 og T3.



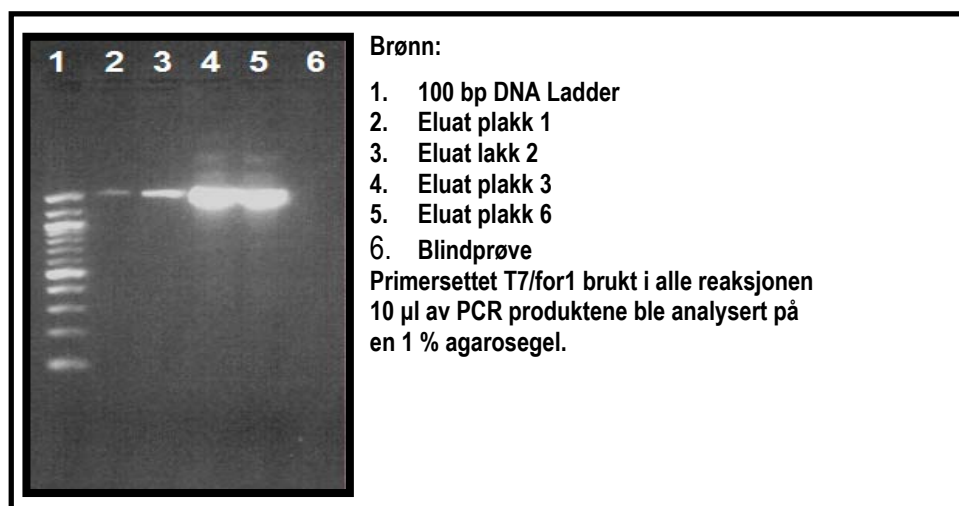
Figur 4.8. PCR på isolerte plakk.

PCR-produktene av primersettet for1 og rew1 på positive plakk ble sendt til sekvensering på MegaBacemaskinen før hele innskuddet ble isolert og sendt til sekvensering.

4.2.4 Isolering av innskudd fra positiv fagklon

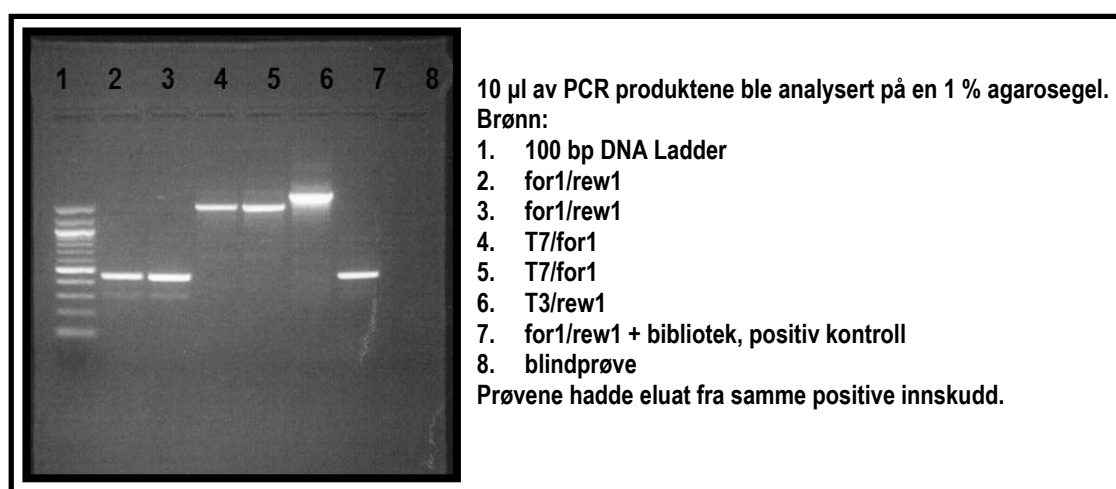
Omkring 1500 bp av *groESL1* var allerede kjent. Sekvensen til *groESL2* var kjent og hadde en størrelse på 2800 bp. Ut i fra dette og ut i fra størrelsen til Hsp60 proteiner, 60 kDa, ble det antatt at *groESL1* var av omtrent samme lengde. Det var derfor ønskelig å finne en fortsettelse på sekvensen på omtrent 1500 bp. Da for1 eller for2 primerene ble brukt til å identifisere fag med riktige fragmenter, og disse primerene bandt ca. 800 bp oppstrøms for den allerede kjente sekvensen, var den ønskelige størrelsen på over 2000 bp. Ved å utføre PCR på positive plakk med kombinasjoner av vektorspesifikke primere med genspesifikke primere og vektorspesifikke primere alene, var det mulig å bestemme retningen av innskuddet i vektoren, størrelsen av innskuddet og beregne om hele *groESL1* operonet var tilstede på den positive klonen. Flere positive plakk ble testet. Ved bruk av primersettet T7 og for1 var det mulig å finne ut om fragmentet var verdt å sekvensere. Var det mindre enn 2000 bp var det ikke av interesse fordi det ikke ville gi noe mer informasjon. Ved bruk av primerene T7 og T3 var ønsket bånd rundt 3000 bp.

De positive plakkene som ga bånd av den ønskede størrelsen, ble videre undersøkt før de ble sendt til sekvensering, for å undersøke om deler av det isolerte fragmentet lå oppstrøms eller nedstrøms for de genspesifikke primerene. Bildet på figur 4.9 viser resultatet fra analyse av 4 plakk. Produktene er oppnådd ved hjelp av primersettet T7 og for1 på eluat fra positive plakk. Alle hadde innskudd på rundt 1500 bp. Dette var kort for å kunne få mer informasjon om nedstrøms sekvensen til *groEL1*.



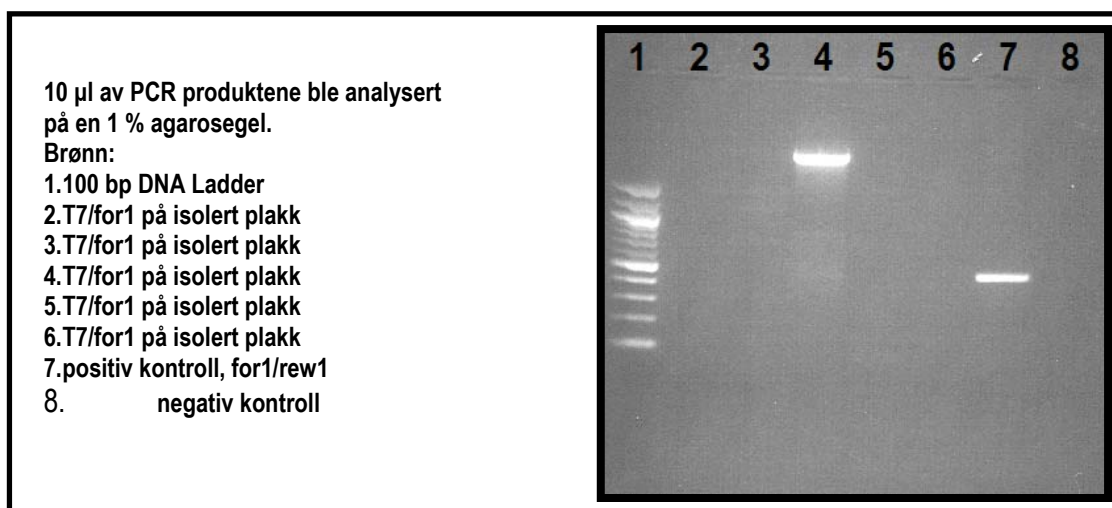
Figur 4.9. PCR med genspesifikk foroverprimer og vektorspesifikk revesprimer på positiv plakk.

Flere innskudd ble isolert og behandlet på samme måte. Noen innskudd var for korte, andre viste forlengelse i gal retning. Noen ga ikke produkt ved bruk av kombinasjonen vektorspesifikke primere med genspesifikke primere. I noen tilfeller ga ikke kombinasjonen av vektorspesifikke primere med genspesifikke primere noe resultat. Dette kan tyde på at den genspesifikke primeren lå helt i enden av innskuddet, noe som i kombinasjon med vektorspesifikk primer ville gi produkter på ca. 100 bp.



Figur 4.10. Analyse av positiv plak

Bildet på figur 4.10 viser at produktene i brønn 2 og 3 hadde kontrollsekvensen på 460 bp i innskuddet. Brønn 4 og 5 viste en sekvens amplifisert fra for1 primeren til vektorspesifikk primer T7. Brønn 6 viste produktet av vektorspesifikk primer T3 til genspesifikk primer rew1. Størrelsen av et av produktene i brønn 4 eller 5 satt sammen med det i brønn 6 gir et fragment på rundt 3000 bp. Dette stemte med størrelsen til innskuddet isolert med T7 og T3 fra positiv plakk. Produktet til T7 og for1 viste et fragment på over 1500 bp. Slike fragmenter ble sendt til sekvensering. Sekvenseringsresultatene er å finne forklart senere i resultatdelen.



Figur 4.11. Analyse av positiv plakk.

Av bildet i figur 4.11 er det i brønn 4 et produkt på rundt 2000 bp. Dette var av ønsket størrelse og ble sendt til sekvensering.

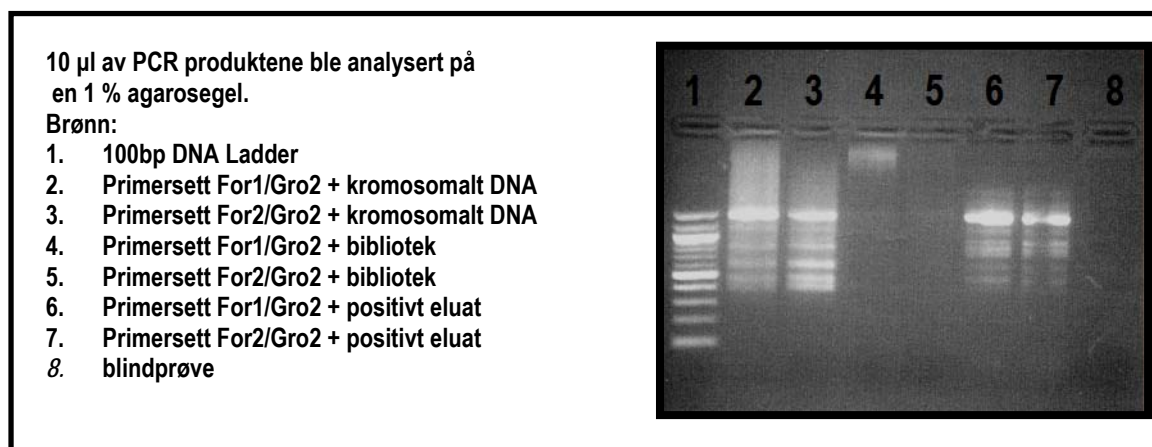
4.3 Søk etter gen ved hjelp av andre primere

På grunn av sekvenseringsresultatene, se punkt 4.4, ble det gjort forsøk på å få resten av sekvensen til operonet ved bruk av en lang sekvenseringsprimer. Denne primeren ble utformet fra et område av den allerede kjente sekvensen og ble brukt på kromosomalt DNA i MegaBasemaskinen. Reaksjonene ble satt opp i samarbeid med avdelingsingeniør Bård Mathisen. Resultatet var dårlige kromatogram med mye bakgrunnsstøy. Det var ikke mulig å lese noen sekvens. Sekvensen var i tillegg ikke lenger enn et par hundre baser, og denne ga ikke noe ny informasjon. Oppsettet ble prøvd flere ganger med ulike reaksjonsbetingelser, men uten resultat.

Det ble også gjort forsøk på å søke etter *groESL1* ved å utforme en primer fra sluttsekvensen til *groESL2*. Denne ble kalt Gro2 og brukt i PCR reaksjone sammen med *groESL1* spesifikk forover primer. Resultatet var bånd på rundt 1500 bp. Noe uspesifikk binding og mindre bånd kan sees i prøvene på figur 4.15.

Av bildet som er tatt av gelen, vist på figur 4.12, kan man se at primersettene som ble brukt ga uspesifikk binding. Primersettene ble brukt på 3 forskjellige templat. Med kromosomalt DNA som templat ble det mer uspesifikk binding enn på bibliotek og positivt eluat som templat. Programmene og reaksjonene ble prøvd ut med ulike parametere og betingelser. Mulig har primeren Gro2 bundet seg til *groESL2*. På biblioteket ble det ikke noe produkt. Gro2 primeren vil ha høyere affinitet til genet det er utformet til, og mest sannsynlig vil ikke de to genene være å finne på samme innskudd. Primersettet kan derfor ikke lage noe produkt. Man kan skimte noen ekstra bånd, dette kan skyldes uspesifikk binding eller enkelttrådig DNA som produkt fra en av primerene. Ved bruk av primersettet på et positivt eluat fra en sektor kom mye av de samme båndene opp som ved bruk på kromosomalt DNA. Det var også færre store fragmenter og mindre uspesifikk binding. Båndene som kom klarest frem og som kunne vært rensset fra gelen, var på størrelsen 1500 bp. Dette var i korteste laget, da en sekvens på over 2000 bp var nødvendig for å få ny informasjon ved bruk av

primer for1 og for2. I tillegg viste gelbildet dårlig isolerte bånd blandet med ekstra bånd og smørje. Reaksjonen lot seg ikke optimalisere videre.



Figur 4.12. PCR med sekvenseringsprimer på kromosomalt DNA fra *C. aurantiacus*

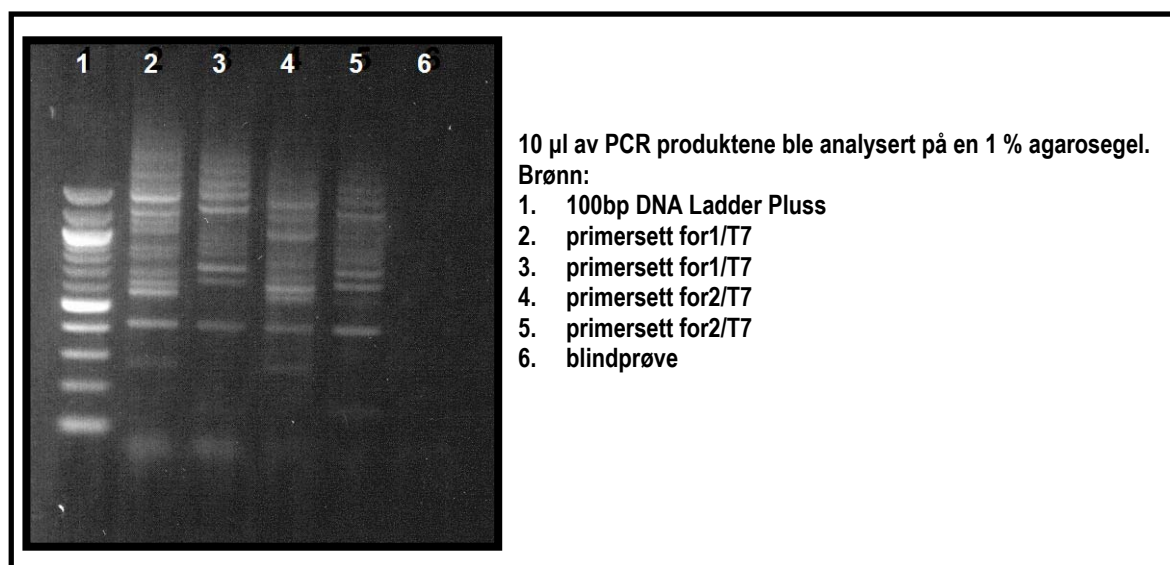
Sekvenser til chaperoniner fra en rekke bakterier, termofile organismer og ikke termofile organismer ble sammenstilt i programmet GenDoc. Chaperoniner er meget konserverte, med høy sekvenslikhet. Både nukleotidsekvens og aminosyresekvens ble sammenstilt. Ved å sammenstille aminosyresekvensene får man en høyere likhetsprosent enn ved nukleotidsammenstilling. I den N-terminale enden hos polypeptidene og den 3'-enden i nukleotidsekvensene er det til tross for dette stor ulikhet. Det var ikke mulig å finne et godt nok konservert område i slutten av sekvensene der en primer kunne utformes.

4.3.1 Bruk av nye primere til "primer-walking"

Da det ved sekvensering viste seg at de isolerte innskuddene fra positive plakk ikke inneholdt det fullstendige genet, ble nye genspesifikke primer utformet i Primer3. Fortsettelsen av sekvensen fra isolerte og sekvenserte positive plakk ble brukt som templat for primerutforming, de ny primerene ble kalt for3, for4, rew3 og rew4. Disse primersettene bandt seg til områder lenger ut i *groEL1* sekvensen og ble brukt på nye søk i genbiblioteket. Ved positive reaksjoner ble søket fortsatt til en plakk var isolert. Yu og Bloems metode ble brukt. Analyse av innskuddet på agarosegel viste hvor stort produktet var og i hvilken retning det strakk seg. På bakgrunn av dette kunne man se om det var verdt å sekvensere, eller om det trolig var kuttet i samme område som de andre sekvensene. Ingen av de isolerte plakkene ga noen informasjon utover det som ble funnet ved bruk av primersettet for1/rew1.

4.3.2 Søk etter gen ved hjelp av T7 og genspesifikke forover primere

Den vektorspesifikke primeren, T7, sammen med en genspesifikk forover primer, for1, ble prøvd ut i en og samme reaksjon direkte på biblioteket. Forventet resultat var bånd av ulike størrelser avhengig av hvor restriksjonsenzymet hadde kuttet i gensekvensen. Reaksjonen ble utført og PCR produktene analysert ved agarose gelelektroforese. Som forventet ble det mange bånd. Bånd av passende størrelse ble rensset ut av gelen ved hjelp av QIAgen gel purification kit. Bildet på figur 4.13 viser at de fleste båndene hadde en størrelse på 1500 bp eller kortere. Bånd over denne størrelsen var diffuse og svake og skyldes sannsynligvis uspesifikk binding av primere til templat.



Figur 4.13 T7 og genspesifikke forover primere på bibliotek som templat.

Fra gelen på bildet, figur 4.13, ble 4 bånd rensset. Disse ble betegnet bånd 1-4. Etter rensing fra gel ble DNAet igjen analysert ved agarose gelelektroforese.

Fordi det kan være problematisk å utføre PCR på PCR produkter ved hjelp av det samme settet med primere, ble det utformet en ny vektorprimer som bandt seg oppstrøms for T7 på vektorarmen. Denne ble brukt sammen med primeren for2 der for1 primeren var brukt og sammen primeren for3 der for2 primeren var brukt. Produktene fra PCR med de rensede båndene som templat, ble sendt til sekvensering hos MWG. Resultatene av sekvenseringen ga ikke noe ny informasjon, og er beskrevet og forklart under avsnittet som omhandler sekvenseringsresultatene, punkt 4.4.

4.4 Aanalyse av sekvenseringsresultatene

Innskudd som inneholdt det antatte *groESL1* operonet og var av ønsket lengde ble sendt til sekvensering. Produkter over 500 bp ble sekvensert av GATC og MWG. Prøven ble sendt som PCR produkt, 30 µl av hver prøve i eppendorfrør. Hver sekvensering tok rundt 14 dager. Resultatene fra GATC kom tilbake som kromatogram og ble analysert og bearbeidet i BioEdit. Sekvensene fra MWG kom tilbake i form av fasta format og som pdf fil av kromatogrammet. Sekvensen ble i første omgang brukt til å søke i databasen til *C. aurantiacus* i NCBI, et BLAST søk.

Sekvensen til *groESL1* som allerede var kjent, Chlo_724, er vist i figur 4.14. Denne sekvensen var på 1430 bp. Da promoterområdet til operonet og *groES1* utgjorde 582 av disse basene, og forventet størrelse til *groEL1* var ca. 2000 bp, manglet sekvensen ca. 1000 bp. Denne sekvensen var allerede kjent (Vanberg, hovedfagsoppgave UIO, 2001) og målet var å finne resten av sekvensen med stopp kodonet, som ville gi et protein på 60 kDa. Alle de første sekvenseringene sluttet litt ut i denne sekvensen og ga ikke noen videre fortsettelse utover det man allerede visste.

```
>gi|22973793|ref|NZ_AAAH01000724.1| Chloroflexus aurantiacus Chlo_724, whole genome shotgun sequence
TCTAGAGGATCCCCCGGCGCTTACTGGCAGGCATTGAGCGTGCCACGCAGCTTGTTAACGACCAGGGTCGCC
AGCGCTTCACCATCAACATCCTCGGCGATGACCAGCAGATCCTTCTCCCGCTGCTCAGCACTGCCTCGAGGATCG
GCAACAGATCTTTGATCGCGCTGATCTTCTGTCGGTGATCAGAATGTATGGCTCATCGAGCACTGCCTCCATGCGG
GCAGAGTCGGTGATGAAGTTTTGCGAGATATAGCCGCGGTGCAATTGCATACCTTCGACCAGCTCGTGCTCGAGAT
GCGTGCTCTTCCCTTCTTCAACCGTTACGACACCATCGCGGCCAATCTTGTCATCACCGTAGCCAGCAGTTCACCG
ACTTCGGCATCCTGGGCCGAGATGGTAGCGACATTCCGAATCTCATCACGGGTTTTGAGCGTAATCGCCTGCTCTTT
GATACGAGCCACGAGCGCTGACCACCCTTATCGAGGCCGCGTTTGAGCAGCATCGGATTGGCACCAGGCTGCAAC
CAGCTTCAAGCCTTCGTTGATGATGGCCTGGCCGAGAACGGTTGCCGTCGTCGTACCGTCACCAGCAACATCATT
GTCTTGACTGCGGCCTGCTTGAGCAACTGCACGCCGAGGTTGGCAAACGGATCTTTCAGTTCAATCTCTTTGGCAAC
GCTCACACCATCGTGGGTTACCGTCGGGGCACCCCACTTCTTGTCGAGCGCAACATTGCGGCCACGTGGGCCAAG
GGTCGTCTTGACAGCCAGAGCCAGCGTATCAATACCGTGTTTCAGCGCTGTACGAGCTTGCTGATCAAAGATCAGT
TGCTTCGCCATACTGCTTACTCCCTTCAATGATACCGAGAATATCCTTCTCCTGGAGAATCAGGTATTCCTCATCGT
CAATCTTAAATCCGTACCGGCATATTTGGCGTAGAGCACGTGCTGACCGACCTTAACACTCATCGGGATCAGTTTG
CCGTCATCAGCACGGCGACCTGGACCCACCGCAATCACTTCACCCTCCATCGGGCGCTCTTTGCTTGCGGTATCGG
GCAGGAAAATGCCGCTTTCGTCTTCTCTTCACGCGGCTTCGGTTTGACTACAACCCGATCATTCAACGGTTTGACG
TGCACAGACGACCTCCTCATTATTCAGGGTTGGCGGTAGTGTCACGCACTACCCACCTTGCCTATGCACGGGTGAC
AGCACACTACCCACCATCGCGTGCCTACTCCGCCGAGGGCAGCGCAACCGCCACACAGATCGGCATCCAC
GCCGTGCCGCGAGGGATCCGACCGCGCTTACCCCGCCGCCATCCGCCACAGCCACCTGTCCGCCCGGCAC
CCCGCAGCGCAAGCGCATCCAGCTCCAGAGCGAAAACCCACCCACAGTCTGAGCGGCC
```

Figur 4.14. Sekvensen til *groESL1* som allerede var kjent og lå ute på NCBI.

Da det viste seg at flere av de sekvenserte positive innskuddene stoppet på samme område, ble sekvensene kuttet *in silico* med restriksjonsenzymet *Sau3AI* som var brukt til å kutte kromosomalt DNA fra *C. aurantiacus* da biblioteket ble konstruert. Programmet som ble brukt til dette var WebCutter. Resultatene fra WebCutter viste at operonet hadde 17 *Sau3AI* kuttsetter, 6 av dem lokalisert i området der sekvenseringen til stadighet stoppet opp. Kuttasetene som ble funnet i dette området er vist på figur 4.15.

Table by Enzyme Name

Enzyme name: Sau3AI No. of cuts: 17 Positions Recognition: gatc
 Sites in sequence 117 145 301 444 596 603 768 1238 1253 1262 1269 1280 1314 1425 1490 1844 2160

	Sau3AI	Sau3AI	Sau3AI	Sau3AI	
gcccgc	catggagg	cagtgctg	atgagccata	cattctgatc	accgacaaga
cgggcgt	acccgctc	acgagctac	cggtatgta	agactagt	ggctgttctt
					tagtcgcg
					ctagttctagac
					base pairs
					1201 to 1275

Sau3AI		Sau3AI	
ttgccgat	cctcgagg	cagtgctg	agcagcgga
aacggctag	gagctcgc	gactcgtc	gcccttctt
			ctctagac
			gaccagtag
			cgcgctc
			tacaactac
			cactt
			base pairs
			1276 to 1350

Figur 4.15. Sau3AI kuttsetter i *groEL1*.

Flere innskudd ble sekvensert og testet. En sekvensering, kalt Plakk2, viste seg å fortsette der Sekvensen fra NCBI sluttet med ca 800 bp. Denne sekvensen er vist på figur 4.16.

>VIII147_Plakk2, 1528 bp, DNA, 27-AUG-2004

```

ctttgatcagcaagctcgtagcgcgtgaaacacggtattgatacgcgtggctctggctgtcaagacgacccttggcccacgtggccgcaa
tggtgcgctcgacaagaagtggggtgccccgacggtaacccacgatggtgtgagcgttgccaaagagattgaactgaagatccgtttg
ccaacctcggcgtgcagttgtcaagcagggccgcagtcgaagacgaatgatgttgctgtgacggtagcagcggcaaccgttctggc
ccaggccatcatcaacgaaggcttgaagctggttcagccgggtgccaatccgatgctgctcaaacgcggcctcgataagggtggtcag
gcgctggtggtcgtatcaagagcagggcgattacgctcaaaaccggtgatgagattcgaatgtcgtaccatctcggcccaggatgc
cgaagtcggtgaactgctggctacgggtatgacaagattggccgcgatggtgtcgtaacggtgaagaagggaagagcagcatctc
gagcagcagctggtcgaaggtatgcaattcgaccgcggctatatctcgccctacttcatcaccgactctgcccgcagtgagcagcggga
gatgagccatacattctgatcaccgacaagaagatcagcgcgatcaaaagatctgttgcgatcctcgaggcagtgctgagcagcggga
agaaggatctgctggtcatcgccgaggatggtgatggtgaagcgtggcgaccctggctgtaacaagctgctggcagcgtcaatgcc
ctggccgtgaaagcggcggcttcggtgatcgccgcaaaagccatgtccaggacattgctatcctaccggtggtaccgtcatctcgga
ggagatcggtcgaaagctcgaaagcgcgaccctgcaagaccttgctgctgcccgcgggtcaaggccgacaagataacaccgttat
cgttgaaggtcacggcgacaagcaggccattcagccggattgccagctcaagcagcagattgaaaccacaacctccgattacga
ccgcgagaagctgcaagagcgggtggcgaagtgtcgggtggtgtcgcgctattaaggttgccgcaccgaccgaaccggccatgaa
agagcgcaaaagctcggttgaagatgcgctcaacgaaccagggtcgggtcgaagagggcatcgtgcccgggtggtggtgtgcctt
gctgaatgcgatccggcgtgataatgtcacgaccagtttgaagaagagcggatggcgctcaatgtcttgcggcgggcgtcgaag
agccactgcgcagctcgccaccaatccgggtgaagatggtcagtggtggtcgagaatgttcgcaacgagcagcgggaagcataaca
acaaccactacggctacgatgtgatgaccggtacctacgtcgatcgaagaattcaaaagcttctcgagagtacttctagagcggccgc
gggcccatcgatttccaccgggtg

```

Figur 4.16. Nukleotidsekvensen til Plakk 2.

Sekvensene ble behandlet med VecScreen, et program distribuert av NCBI, der man kan søke etter vektorkontaminering i sekvenser. Resultatene viste kontaminering i området 1454-1528 bp, de siste 75 basene i sekvensen. Dette var som forventet da T7 primer var brukt for å isolere innskuddet. T7 bindes litt innenfor restriksjonsenzymets kuttsete. Disse basene ble fjernet fra sekvensen. Sekvensen ble videre behandlet med Transeq, et program fra EMBL-EBI. Dette programmet brukes til å utlede proteinsekvenser fra nukleotidsekvenser. Primeren for1 var brukt til å isolere innskuddet. Da denne begynte litt ut i *groEL1*, manglet sekvensen begynnelsen av genet.

>groESL1

```
GGGCCGCTCAGACTGGGGTGGGGTTTTCTGCTCTGGAGCTGGATGCGCTTGC GCGTGC GGGGTGCGGGG
CGGACAGGTGGCTGTGGGCGGATGGGCGGGCGGGGTGAAGCGGCGGTTCGGATCCCTCGCGGCACGGC
GTGGGATGCCGATCTGTGTGGCGGTTGCCGTGCGCTCCGGGCGGAGTGCACACGCGATGGTGGGTGAGT
GTGCTGTCAGCCGTGCATAGGCAAGGTGGGTAGTGGCTGACACTACCGCCAAACCTGAATAATGAGGAGG
TCGTCTGTGCACGTCAAACCGTTGAATGATCGGGTTGTAGTCAAACCGAAGCCGCGTGAAGAGAAGACG
AAAGGCGGCATTTTCTGCCCCGATACCGCAAGCAAAGAGCGCCCCGATGGAGGGTGAAGTGATTGCGGT
GGGTCCAGGTGCGCGTGCTGATGACGGCAAACTGATCCCGATGAGTGTTAAGGTGCGTCAGCACGTGCT
CTACGCCAAATATGCCGGTACGGAATTTAAGATTGACGATGAGGAATACCTGATTCTCCAGGAGAAGGAT
ATTCTCGGTATCATTGAAGGGAGTAAAGCAGTATGGCGAAGCAACTGATCTTTGATCAGCAAGCTCGTAC
AGCGCTGAAACACGGTATTGATACGCTGGCTCTGGCTGTCAAGACGACCCCTTGCCCCACGTGGCCGCAA
TGTTGCGCTCGACAAGAAGTGGGGTGGCCCGACGGTAACCCACGATGGTGTGAGCGTTGCCAAAGAGAT
TGAAGTGAAAGATCCGTTTGCCAACCTCGGCGTGCAGTTGCTCAAGCAGGCCGCGAGTCAAGACGAATGA
TGTGTGCTGGTGACGGTACGACGACGGCAACCGTTCTGGCCGAGGCCATCATCAACGAAGGCTGAAGCT
GGTTGCAGCCGGTGCCAATCCGATGCTGCTCAAACGCGGCCCTCGATAAGGGTGGTCAGGCGCTGGTGG
CTCGTATCAAAGAGCAGGCGATTACGCTCAAACCCGCTGATGAGATTGGAATGTCGCTACCATCTCGGC
CCAGGATGCCGAAGTCGGTGAACTGCTGGCTACGGTGATGGACAAGATTGGCCGCGATGGTGTGCTAAC
GGTTGAAGAAGGGAAGAGCACGCATCTCGAGCACGAGCTGGTCGAAGGTATGCAATTCGACCGCGGCT
ATATCTCGCCCTACTTCATCACCGACTCTGCCCGCATGGAGGCAGTGCTCGATGAGCCATACATTCTGAT
CACCGACAAGAAGATCAGCGCGATCAAAGATCTGTTGCCGATCCTCGAGGCAGTGCTGAGCAGCGGGA
AGAAGGATCTGCTGGTCATCGCCGAGGATGTTGATGGTGAAGCGCTGGCGACCCTGGTCGTTAAACAAGC
TGCGTGGCACGCTCAATGCCCTGGCCGTGAAAGCGCCCGGCTTCGGTGATCGGCGCAAAGCCATGCTCC
AGGACATTGCTATCCTACCGGTGGTACCGTCACTCGGAGGAGATCGGTGCGAAGCTCGAAAGCGCGA
CCCTGCAAGACCTTGGTCTGTCGCCCGCGGGTCAAGGCCGACAAAGATAACACCGTTATCGTTGAAGGTC
ACGGCGACAAGCAGGCCATTGAGGCCCGGATTGCCAGCTCAAGCAGCAGATTGAAACCAACCTCC
GATTACGACCCGCGAGAAGCTGCAAGAGCGGGTGGCGAAGTTGTGCGGTGGTGTGCGCGTCATTAAGGT
GGCGCACCGACCGAACCGGCCATGAAAGAGCGCAAAGCTCGCGTTGAAGATGCGCTCAACGCAACCGAG
GGCTGCGGTGCAAGAGGGCATCGTGCCCGTGGTGGTGTGTCCTGCTGAATGCGATCCCGGCGCTTGA
TAATGTCACGACCCAGTTTGAAGAAGAGCGGATGGCGCTCAATGTCTTGC GGCGGGCGCTCGAAGAGCC
ACTGCGCCAGCTCGCCACCAATGCCGGTGAAGATGGTTCAAGTGGTGGTGGTGAAGATGTTGCGCAACGAGCA
GCGGAAGCATAACAACAACCACTACGGCTACGATGTGATGACCGGTACCTACGTCGATTCTGATGCAGT
CCGGCATTATCGACCCGGCGAAGGTAGTGCGCTCCGCTTTTCTCGAAAAATGCGAATCAGCGTTGCCG
GGATGGTGCTGACACCGAGGCCCTGATCGTGAATGCACCCGAACCGAAGAAGAAGAACGGTACGCC
ACCGATGCGCGACGATGACTTCTAA
```

Figur 4.18. Nukleotidsekvensen til *groESL1*.

Sekvensen til genet var på 2228bp. Sekvensen ble igjen oversatt til proteinsekvens og molekylvekten beregnet. Dette er vist i figur 4.19.

>virt|VIRT15615|VIRT_15615 Translation of nucleotide sequence generated on ExPASy on 26-Jan-2005
by ti200720a080-1983.bb.online.no

```
MAKQLIFDQQARTALKHGIDTLALAVKTTLGPGRNVALDKKWGAPTVDHGVSVAKEIELKDPFANLGVQ
LLKQAAVKTNVDAGDGTATVLAQAIINEGLKLVAAAGANPMLLKRGLDKGGQALVARIKEQAITLKTRDEIR
NVATISAQDAEVGELLATVMDKIGRDGVVTVEEGKSTHLEHELVEGMQFDRGYISPYFITDSARMEAVLDEPY
ILITDKKISAIKDLLPILEAVLSSGKKDLLVIAEDVDGEALATLVVNKLRTLNALAVKAPGFGDRRKAMLQDIA
ILTGGTVISEEIGRKLESATLQDLGRARRVKADKDNVTIVEGHGDKQAIQARIAQLKQIETTTSDYDREKLQE
RVAKLSGGVAVIKVGAPTEPAMKERKARVEDALNATRAAVEEGIVPGGGVALLNAIPALDNVTTQFEERM
ALNVLRRALEEPLRQLATNAGEDGSVVVENVRNEQRKHNNHYGYDVMGTGYVDSDAVRHYRPGEGSALR
ESRKNANQRCRDGADHRGPDRRMHPNRRRTVRHRCRTMTS
```

Figur 4.19. Aminosyresekvensen til *groEL1*

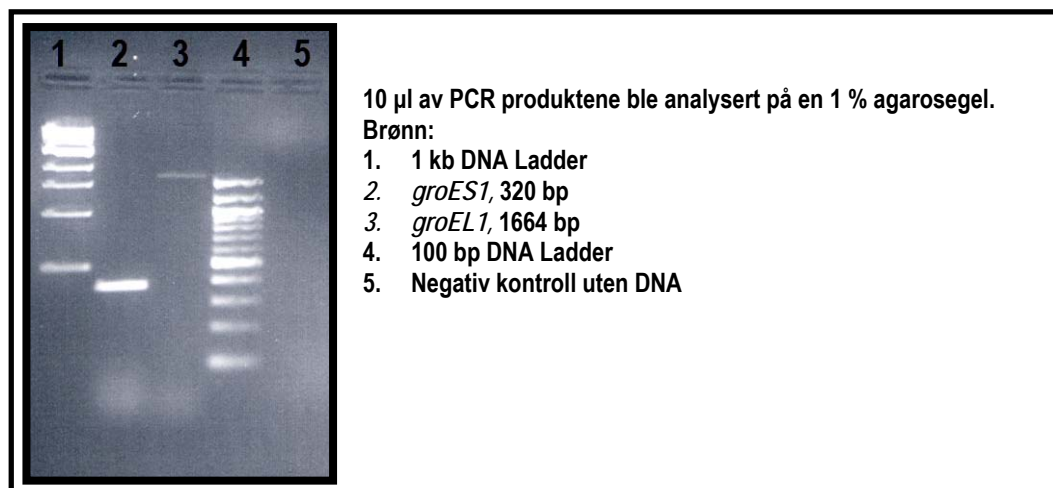
Molekylvekten på dette proteinet var 59,7 kDa. Ved søk med dette i *C. aurantiacus* protein database på BLAST fikk man opp chaperoninet GroEL. Resultatet viste 97% sekvenslikhet. Deler av sekvenssammenstillingen er vist i figur 4.20. Sekvensen til Plakk2, Chlo01_574 og Chlo_724 ble sammenstilt i Multalin. Sammenstillingen er å finne i Appendiks.

> ref ZP_00359398.1 COG0459: Chaperonin GroEL (HSP60 family) [Chloroflexus aurantiacus] Length = 282	
Score = 514 bits (1323), Expect = e-147 Identities = 272/279 (97%), Positives = 272/279 (97%)	
Query: 1	MAKQLIFDQQARTALKHGIDTLALAVKTTLGPRGRNVALDKKWGAPTVTHDGVSVAKEIE 60
Sbjct: 1	MAKQLIFDQQARTALKHGIDTLALAVKTTLGPRGRNVALDKKWGAPTVTHDGVSVAKEIE 60
Query: 61	LKDPFANLGVQLLKQAAVKTNDVAGDGTATVLAQAIINEGLKLVAAGANPMLLKRGD 120
Sbjct: 61	LKDPFANLGVQLLKQAAVKTNDVAGDGTATVLAQAIINEGLKLVAAGANPMLLKRGD 120
Query: 121	KGGQALVARIKEQAITLKTRDEIRNVATISAQDAEVGELLATVMDKIGRDGVVTVVEEGKS 180
Sbjct: 121	KGGQALVARIKEQAITLKTRDEIRNVATISAQDAEVGELLATVMDKIGRDGVVTVVEEGKS 180
Query: 181	THLEHELVEGMQFDRGYISPYFITDSARMEAVLDEPYILITDKKISAIKDLLPILEAVLS 240
Sbjct: 181	THLEHELVEGMQFDRGYISQNFITDSARMEAVLDEPYILITDKKISAIKDLLPILEAVLS 240
Query: 241	SGKKDLLVIAEDVDGEALATLVVNKLGRGTNLALAVKAPG 279
Sbjct: 241	SGKKDLLVIAEDVDGEALATLVVNKLGRGTNLNA K G 279

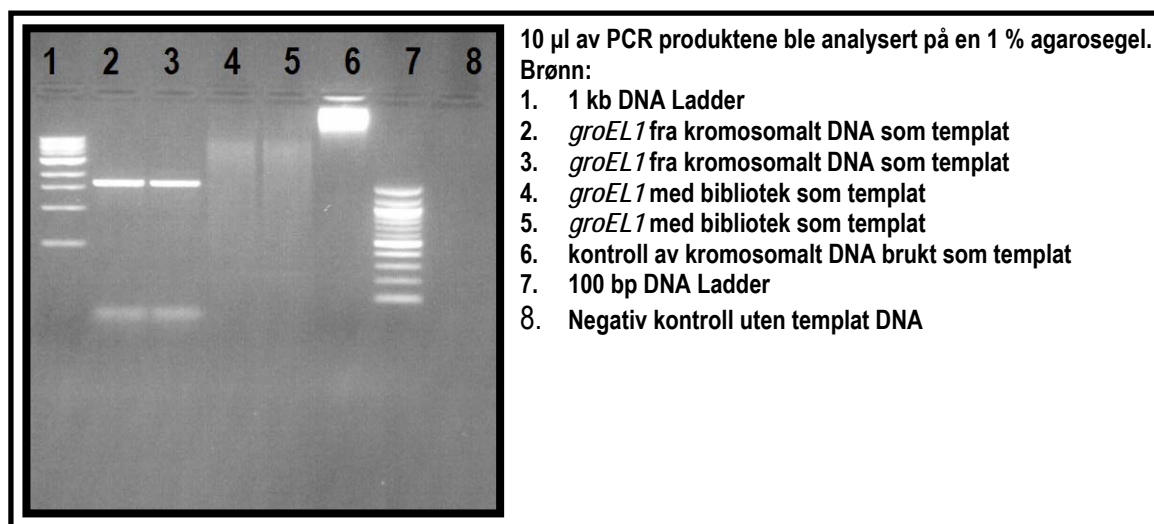
Figur 4.20. Sekvenssammenstilling av utledet sekvens og sekvensen til *groEL1* fra NCBI.

4.5 Isolering av *groES1* OG *groEL1*

Da hele operonet var sekvensert, ble det utformet to sett primere, det ene GroES1F/GroES1R for å isolere *groES1* og det andre GroEL1F/GroEL1R for å isolere *groEL1*. Da primersettet til *groEL1* var basert delvis på *in silico* forsøk var størrelsen av det isolerte genet svært interessant ved første PCR reaksjon. Et fragment med uriktig størrelse ville tyde på at sekvensen ikke stemte med hva som var å finne i organismen. Forventede bånd for *groES1* var 320 bp og for *groEL1* 1664 bp. Resultatet av PCR reaksjonen er vist på gelbildet på figur 4.21.

Figur 4.21 Isolert *groES1* og *groEL1*

Det viste seg at primersettet utformet til *groES1* ga gode resultater i PCR ved bruk av både bibliotek og kromosomalt DNA som templat. Primersettet til *groEL1* fungerte bare på kromosomalt DNA. Dette styrket teorien om at *groEL1* ikke eksisterer i sin helhet i genbiblioteket på grunn av mange *Sau3AI* kutteseter i slutten av denne nukleotidesequensen. Bildet av gelen, figur 4.22, viser resultatet av primersettet GroEL1F/GroEL1R. I brønn 2 og 3 er kromosomalt DNA fra *C. aurantiacus* brukt som templat. Man kan se et klart bånd på 1664 bp i hver av brønnene. Det svake båndet nederst er primer-dimer. I brønn 4 og 5 er bibliotek brukt som templat med det samme primersettet. Her er det ikke noe klare bånd, men noen svake bånd som kan tyde på uspesifikk, men ikke sterk binding av primerene til templat. Da det var en stund siden det isolerte DNAet var tatt i bruk, ble dette undersøkt på gel, brønn 6. DNAet fungerte som det skulle med primersettene og viste ingen tegn for degradering i form av smørje på gelbildet.

Figur 4.22. Isolering av *groEL1*

Det ble forsøkt å utforme et primersett til isolering av hele operonet, men sekvensen oppstrøms for operonet viste seg å være for kort til å kunne utforme primere som ville være spesifikke nok.

De to isolerte genene ble sendt til sekvensering for å bekrefte at riktige gener var isolert og for å se at det ikke inneholdt noen mutasjoner. Fragmentet *groES1* ble sendt til sekvensering på Megabasemaskinen. Sekvensen viste seg å være korrekt. På grunn av lengden til *groEL1* ble denne sendt til MWG. Sekvenseringen viste seg å være svært problematisk. Bare et par hundre baser ble sekvensert i hver retning. Etter flere mislykkede sekvenseringer ble det antatt at sekvensen til *groEL1* inneholdt tredimensjonale strukturer som vanskelig lot seg sekvensere. Sekvensen ble derfor analysert i programmet mfold, et windowsbasert program som brukes til å folde enkelttrådig DNA og RNA sekvenser. Enkelttrådig DNA vil ha de samme foldingsbetingelsene som RNA. Sekvensering av DNA skjer ved 44 °C og DNAet er enkelttrådig. Det finnes flere ulike foldingsstrukturer, blant annet hårnålsstrukturer og pseudoknuder. Er bindingene sterke (G/C-rik) og mange vil de kunne ødelegge sekvenseringen. Folding av enkelttrådig DNA er beskrevet i innledningen, punkt 1.5.

Sekvensen til *groEL1* ble analysert i programmet og beregningene gjort. Ut fra strukturene som ble konstruert kunne man se at sekvensen inneholdt flere sekundære strukturer. Sekvensen til *groEL2* ble også satt inn i programmet for å kunne sammenligne de to genene. Sekvensene ble både satt inn i sin helhet, og også delt på midten slik at starten og slutten ble analysert hver for seg. Dette fordi *groEL1* viste seg å være vanskeligst å sekvensere fra midten av genet og ut sekvensen.

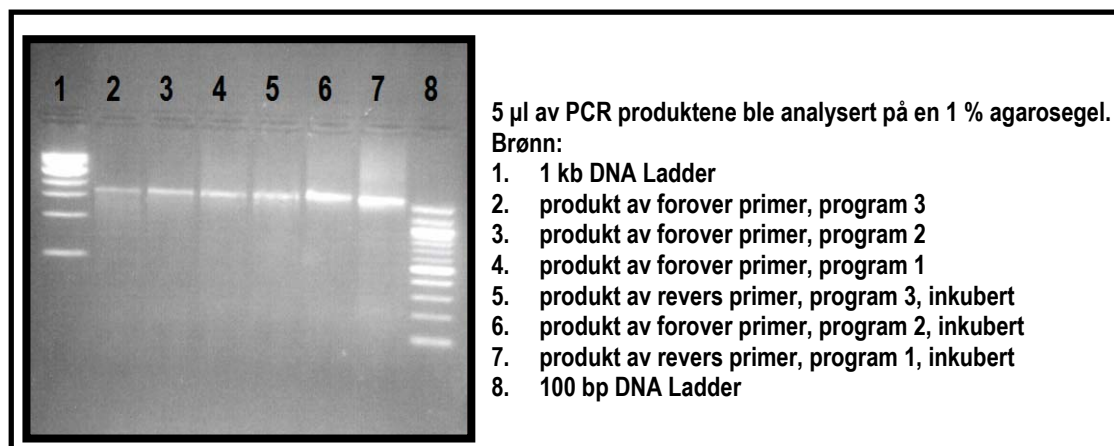
Resultatene til mfold viste de mest sannsynlige foldingsegenskapene til sekvensene sammen med dinukleotidhyppigheten til basene. I området det var problemer med sekvenseringen viste *groEL1* en mye høyere dinukleotidhyppighet av baseparet GC enn i områdene sekvenseringen ikke var problematisk. Hyppighet av GC var høyere her enn i det tilsvarende området til *groEL2*.

Resultatene av det som var sekvensert ga treff i BLAST på *groEL1* fra *C. aurantiacus* og viste seg å være identisk med *groEL1* da sekvensene ble sammenstilt i Multalin. Den ble derfor brukt videre i ekspresjon.

4.6 Sekundære strukturer i DNA

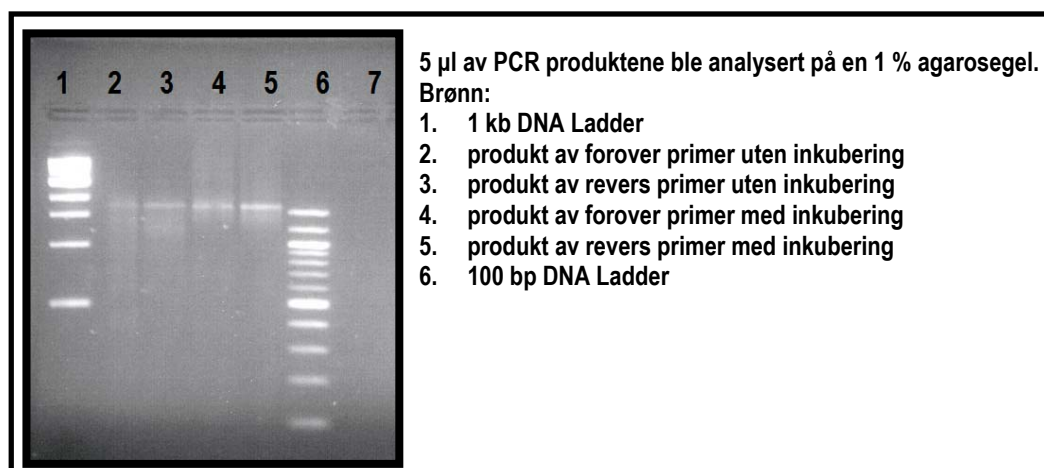
Da det viste seg at *groEL1* inneholdt sekundære strukturer som gjorde sekvenseringen vanskelig, ble det satt opp 3 alternative PCR programmer for å prøve å forhindre dannelsen av de sekundære strukturene i enkelttrådig DNA. Programmene som ble brukt er beskrevet i punkt 3.12. Produktene ble sekvensert på MegaBace maskinen. Programmet der det denaturerende trinnet var på 95 °C i 30 sekunder og elongeringstrinnet på 60 °C i 240 sekunder (program 3) viste seg å fungere best med revers primer. Produktet som ble produsert under disse betingelsene ga en sekvens på omtrent 1000 baser, der ca. 900 baser stemte overens med *groEL1*. Programmet der det denaturerende trinnet var på 95 °C i 30 sekunder og elongeringstrinnet på 60 °C i 120 sekunder (program 2) viste seg å fungere best på forover primer. Produktet som ble produsert under disse betingelsene ga en sekvens på omtrent 500 baser, der ca. 250 baser stemte overens med *groEL1*.

PCR produktene fra hver reaksjon ble delt i to. Den ene delen ble analysert direkte på agarosegel, mens den andre ble inkubert ved 44 °C i 4 - 6 timer før den ble analysert. Dette siste for å simulere behandlingen av DNA i sekvenseringsmaskinen.



Figur 4.23 Produkt av *groEL1* fra alternative PCR sekvenseringsprogram

Av bildet på figur 4.23 ser man et tydelig bånd med størrelse på omkring 1600 bp i alle brønnene. Dette er resten av *groEL1* som ble brukt som templat i reaksjonene. Det ble mer smørje ved bruk av program 1 enn ved bruk av program 3. Smørjen kan være enkelttrådig DNA som har bundet etidiumbromid i områder med sekundære strukturer. Det ble også kjørt to reaksjoner der programmene var de samme som de som vanligvis brukes for sekvensering av PCR produkter. Også disse ble delt i to, med inkubering av den ene delen på 44 °C. Produktene ble analysert på agarosegel.



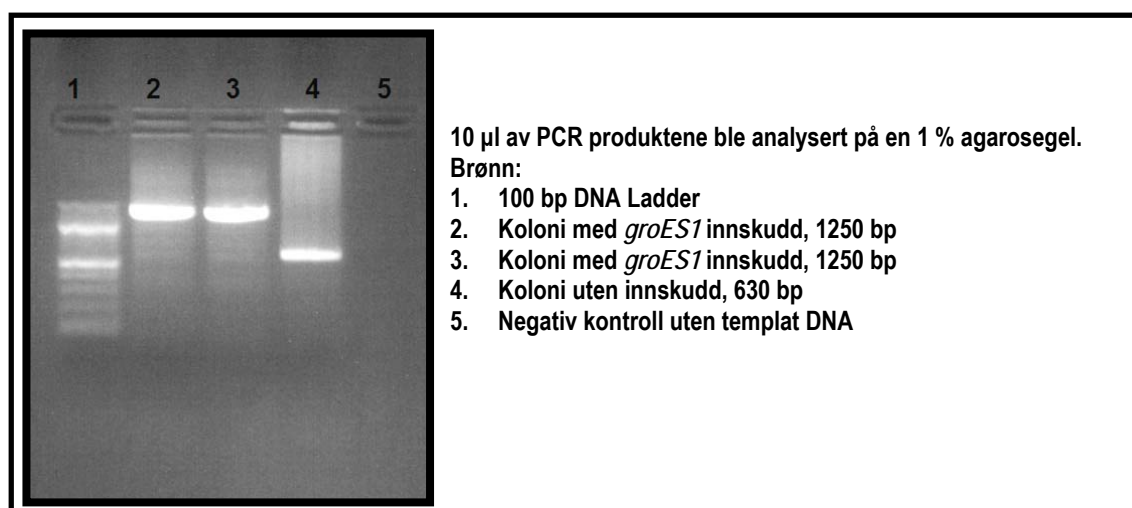
Figur 4.24 Produkt av *groEL1* fra vanlig PCR sekvenseringsprogram.

Båndet på 1600 bp er også synlig på gelbildet som er vist på figur 4.24. Det ble mindre smørje ved bruk av standard oppsettet for sekvensering, enn ved bruk av programmene som var konstruert for vanskelige templat. Sekvenseringsresultatene ga derimot dårligere kromatogrammer.

4.7 Transformer og ekspressjon av *groES1* OG *groEL1*

Det ble utformet to nye forover primere til hver av de to genene, *groES1* OG *groEL1*, til bruk i TOPO kloningsreaksjoner. Disse primerene hadde CACC-overheng før sekvensen fortsatte med ATG, startkodonet til genene. Disse ble brukt på kromosomalt DNA isolert fra *C. aurantiacus*. CACC-overhenget var nødvendig for å ligere genene i TOPO vektoren. PCR reaksjonen ble satt opp med TOPO-primerene og transformeringen av *groEL1* og *groES1* utført som beskrevet i manualen til pET 102/D TOPO settet.

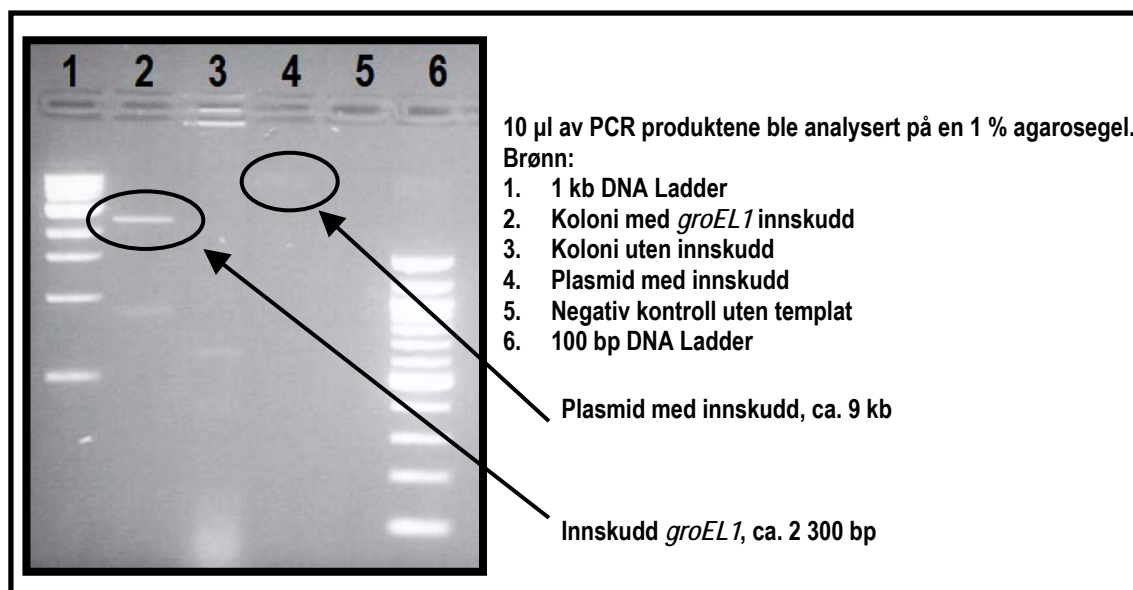
Da pET 101/D TOPO og pET 102/D-TOPO vektorene ikke inneholder noen form for seleksjonsmarkør for om vektoren har tatt opp innskuddet eller ikke, var det nødvendig å teste koloniene for innskudd før vertscellene ble dyrket og plasmider rensset fra cellene. Koloniene ble testet for innskudd i PCR reaksjon med primersettet T7/T7revers, som er de spesifikke primerene til vektoren. I vektoren som ble brukt, pET 102/D-TOPO, festes T7 og T7 revers primerene henholdsvis 466 bp oppstrøms og 164 bp nedstrøms for kloningssetet, slik at når PCR utføres på dette plasmidet var resultatet produkter som var 630 bp lenger enn genene satt inn i vektoren. *GroES1* og *groEL1* ville derfor vises som bånd på henholdsvis 950 bp og 2294 bp. Produktene ble analysert på agarosegel. På figur 4.25 er det vist et bilde av produktene fra PCR på 3 kolonier transformert med *groES1*.



Figur 4.25 PCR på kolonier med og uten *groES1* innskudd.

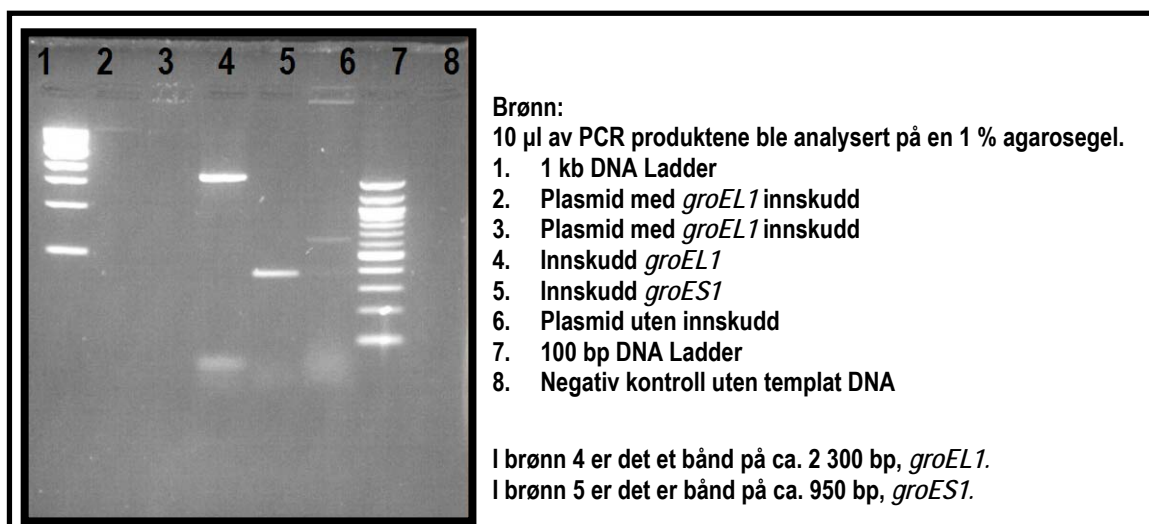
I brønn 4 vises et resultat av primersettet på vektor uten innskudd, i brønn 2 og 3 er PCR kjørt på plasmidet med innsatt *groES1*. *groES1* ble i dette tilfellet isolert med *rew1* primeren som ga et 300 bp større produkt enn ved bruk av *groES1R* primeren.

På bildet på figur 4.26 er det utført PCR på 2 kolonier transformert med *groEL1*. I brønn 2 vises resultatet av primersettet på vektor med innskudd (2200 bp), i brønn 3 vises PCR på koloni uten innskudd (630 bp). I brønn 4 kan plasmidet med innskudd (rundt 9 kb) skimtes, og i brønn 5 er PCR kjørt uten templat, negativ kontroll.



Figur 4.26 PCR på kolonier med *groEL1* innskudd.

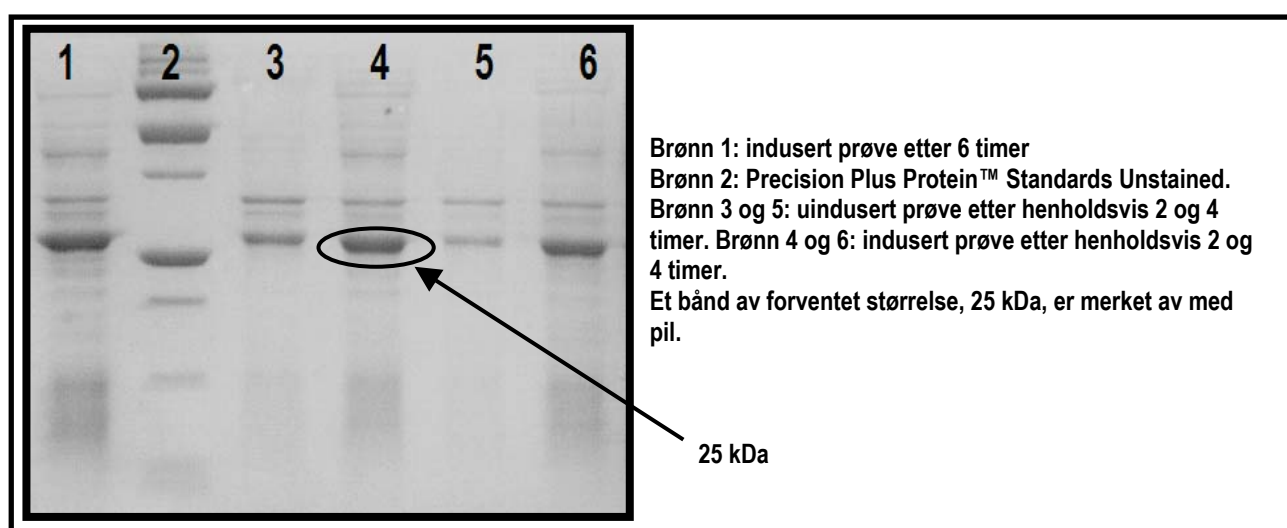
Plasmidene ble rensert som beskrevet i manualen til settet. For å forsikre om at innskuddet ikke var tapt på veien ble det utført en ny PCR med de vektorspesifikke primerene på plasmid. Resultatet er vist på gelbildet på figur 4.27. I brønn 2 og 3 kan man skimte plasmidene. Størrelsen på plasmidet er 6 315 bp uten innskudd. Ut i fra størrelsesstandard kan man se at plasmidene inneholdt innskudd. PCR produktene i brønn 4 og 5 viser de to genene. Genspesifikke primere ble brukt på plasmidene. I brønn 6 er det vist PCR på plasmid uten innskudd og brønn 8 er negativ kontroll uten templat DNA.



Figur 4.27 PCR på rensede plasmider

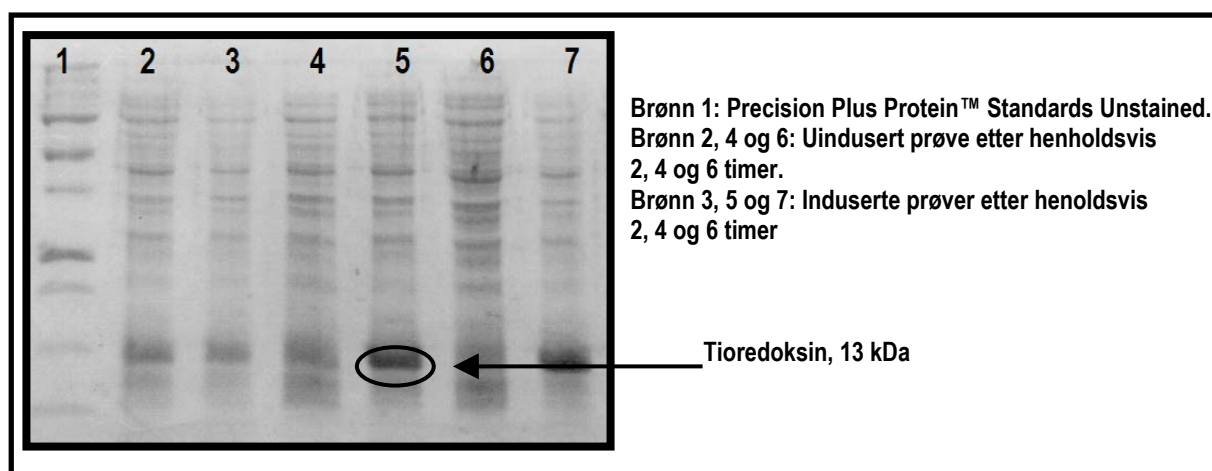
Plasmidene med innskudd ble også sendt til sekvensering på Megabace maskinen for å forsikre om at innskuddet var satt riktig inn i vektoren, med tanke på startkodon og leserramme. Ekspresjonen av de to genene ble utført som beskrevet i manualen med pET 101/D-TOPO og pET 102/D-TOPO vektorene. Vertscellene var BL21star celler. Resultatene ble analysert på SDS-PAGE.

Fordi vanlig SDS-PAGE kan gi dårlig oppløsning av proteiner under 15 kDa og fordi *groES2* fra *C. aurantiacus* var uttrykt med hell som fusjonsprotein (Benestad, hovedfagsoppgave UIO, 2005), ble genet *groES1* forsøkt uttrykt som fusjonsprotein, fusert med tioredoksin. Vektoren pET 102/D TOPO ble brukt til dette. Molekylvekten til tioredoksin er 13 kDa og for GroES1 11,5 kDa. Et protein på rundt 24,5 kDa var derfor forventet. Av bildet på figur 4.28 kan man se produktet av *groES1* med tioredoksin som fusjonsprotein, omkring 25 kDa.

Figur 4.28 Ekspresjon av *groES1*

Det ble ikke observert noen ekspresjon av *groEL1* som fusjonsprotein ved bruk av pET 102/ D TOPO vektoren. Ekspresjonsproduktet av *groEL1* med fusjonsproteinet ville hatt en størrelse på 73 kDa. Båndet nederst på gelen viser ekspresjon av tioredoksin, 13 kDa, noe som viser at induksjonsforsøket gikk som det skulle. Dette er vist på bildet av gelen, figur 4.29

Det ble forsøkt å uttrykke *groEL1* ved hjelp av pET 101/ D TOPO vektoren, som er en vektor der genet blir uttrykt uten fusjonsprotein. Det ble ikke observert noen ekspresjon av genet da celleekstraktet fra induerte BL21 celler ble analysert på SDS-PAGE. Plasmidene som cellene var transformert med var analysert ved PCR og viste innskudd. Innskuddene var sekvensert for å forsikre om at de var satt inn i vektoren korrekt og i leserramme.



Figur 4.29 Ekspresjon av *groEL1* i BL21 Star celler. Bildet viser uttrykkelse av fusjonsproteinet tioredoksin.

Det anbefalt av produsenten til pET vektorene, Invitrogen, å uttrykke gener med vanskelige sekundære strukturer i 30 °C over 12 timer. Dette for at BL21 cellenes transkripsjonsmaskineri skulle få tid til å lese gjennom genene og ikke "hoppe av" i områder med sekundære strukturer. Det ble forsøkt med ulike konsentrasjoner av IPTG og ulike konsentrasjoner plasmid. BL21 cellene ble inkubert med og uten glukose tilsatt i mediet, ved 37 °C og 30 °C, over 6 og 12 timer. Induserte celler ble også satt til inkubering ved romtemperatur over 24 timer.

Da dette ikke ga noen ekspresjon av *groEL1*, ble det sett etter om genet inneholdt kodoner som sjeldent brukes av *E. coli*. Sjeldne kodoner kan forårsake ingen ekspresjon eller for tidlig terminering av heterologe gener. Det ble ved hjelp av programmet Rare Codon Calculator funnet til sammen 14 sjeldne kodoner i *groEL1*, disse er vist i figur 4.30.

```

>groEL1      del 1
atg gcg aag caa ctg atc ttt gat cag caa gct cgt aca gcg ctg aaa cac ggt att gat acg ctg gct ctg gct gtc aag
acg acc ctt ggc cca cgt ggc cgc aat gtt gcg ctc gac aag aag tgg ggt gcc ccg acg gta acc cac gat ggt gtg
agc gtt gcc aaa gag att gaa ctg aaa gat ccg ttt gcc aac ctc ggc gtg cag ttg ctc aag cag gcc gca gtc aag
acg aat gat gtt gct ggt gac ggt acg acg acg gca acc gtt ctg gcc cag gcc atc atc aac gaa ggc ttg aag ctg
gtt gca gcc ggt gcc aat ccg atg ctg ctc aaa cgc ggc ctc gat aag ggt ggt cag gcg ctg gtg gct cgt atc aaa
gag cag gcg att acg ctc aaa acc cgt gat gag att cgg aat gtc gct acc atc tcg gcc cag gat gcc gaa gtc ggt
gaa ctg ctg gct acg gtg atg gac aag att ggc cgc gat ggt gtc gta acg gtt gaa gaa ggg aag agc acg cat ctc
gag cac gag ctg gtc gaa ggt

      del 2
atg caa ttc gac cgc ggc tat atc tcg CCC tac ttc atc acc gac tct gcc cgc atg gag gca gtg ctc gat gag cca tac
att ctg atc acc gac aag aag atc agc gcg atc aaa gat ctg ttg ccg atc ctc gag gca gtg ctg agc agc ggg aag
aag gat ctg ctg gtc atc gcc gag gat gtt gat ggt gaa gcg ctg gcg acc ctg gtc gtt aac aag ctg cgt ggc acg ctc
aat gcc ctg gcc gtg aaa gcg

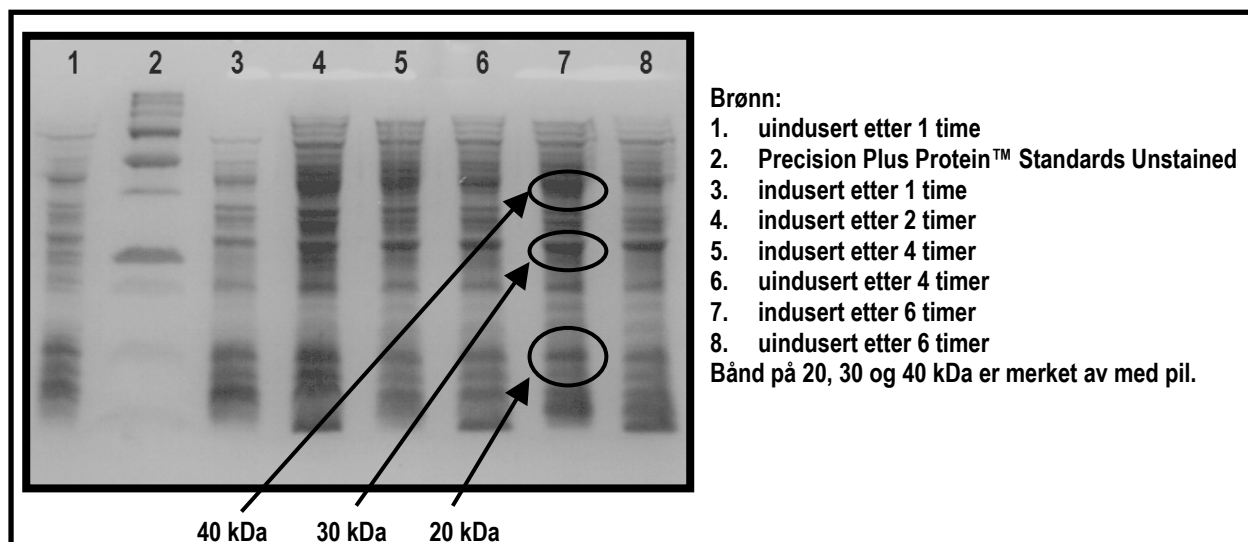
      del 3
CCC ggc ttc ggt gat cgg cgc aaa gcc atg ctc cag gac att gct atc ctc acc ggt ggt acc gtc atc tcg gag gag
atc ggt cgc aag ctc gaa agc gcg acc ctg caa gac ctt ggt cgt gcc cgc cgg gtc aag gcc gac aaa gat aac acc
gtt atc gtt gaa ggt cac ggc gac aag cag gcc att cag gcc cgg att gcc cag ctc aag cag cag att gaa acc aca
acc tcc gat tac gac cgc gag aag ctg caa gag cgg gtg gcg aag ttg tcg ggt ggt gtc gcc gtc att aag gtt ggc
gca ccg acc gaa ccg gcc atg aaa gag cgc aaa gct cgc gtt gaa gat gcg ctc

Del 4
aac gca acc AGG gct gcg gtc gaa gag ggc atc gtg CCC ggt ggt ggt gtt gcc ctg ctg aat gcg atc ccg gcg ctt
gat aat gtc acg acc cag ttt gaa gaa gag cgg atg gcg ctc aat gtc ttg cgg cgg gcg ctc gaa gag cca ctg cgc
cag ctc gcc acc aat gcc ggt gaa gat ggt tca gtg gtg gtc gag aat gtt cgc aac gag cag cgg aag cat aac aac
aac cac tac ggc tac gat gtg atg acc ggt acc tac gtc gat tct gat gca gtc cgg cat tat CGA CCC ggc gaa ggt
agt gcg ctc cgc ttt tct CGA aaa aat gcg aat cag cgt tgc cgg gat ggt gct gac cac CGA ggc cct gat cgt CGA
atg cac ccg aac CGA AGA AGA AGA acg gta cgc cac CGA tgc cgg acg atg act tct

```

Figur 4.30. *GroEL1* analysert i programmet Rare Codon Calculator.

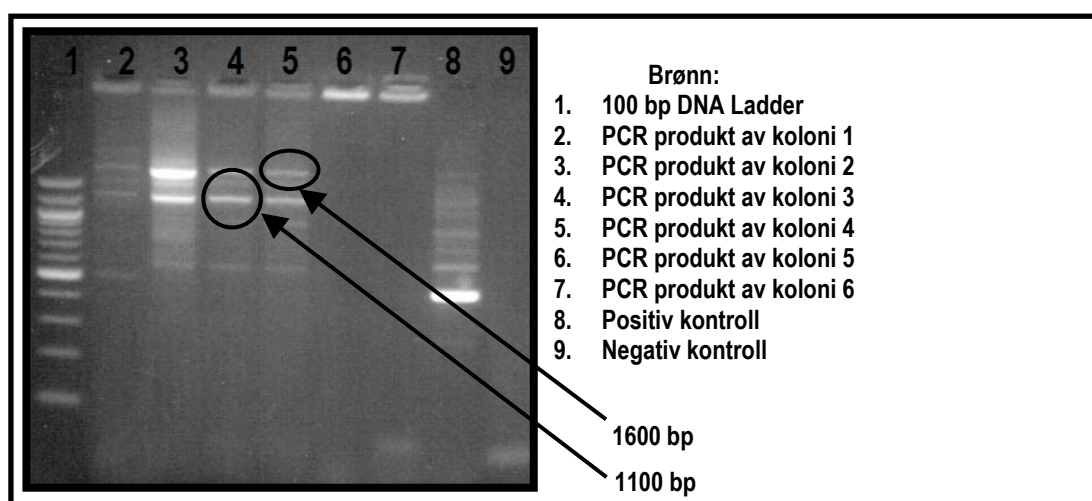
Kodonene merket blått koder for prolin, kodonene merket rødt for arginin. De aller fleste befant seg på den C-terminale enden av den kodende sekvensen. Genet ble delt opp slik som vist over, og delene translatert til aminosyresekvenser ved hjelp av programmet ExPASy-Transeq. Dette programmet beregnet også molekylvekten til proteinet. Dersom translasjonen termineres på grunn av sjeldne kodoner, slik at kun den øverste delen translateres, vil produktet bli et protein på rundt 20 kDa. Proteinproduktet vil være rundt 30 kDa dersom translasjonen fortsetter ut del to. Dersom translasjonen fortsetter ut del tre, vil det gi et proteinprodukt på rundt 43 kDa. Hele *groEL1* vil, som fortalt tidligere, gi et produkt på ca 60 kDa.



Figur 4.31 Ekspresjon av *groEL1* i BL21 Star celler. Bildet viser for tidlig terminering av translasjon.

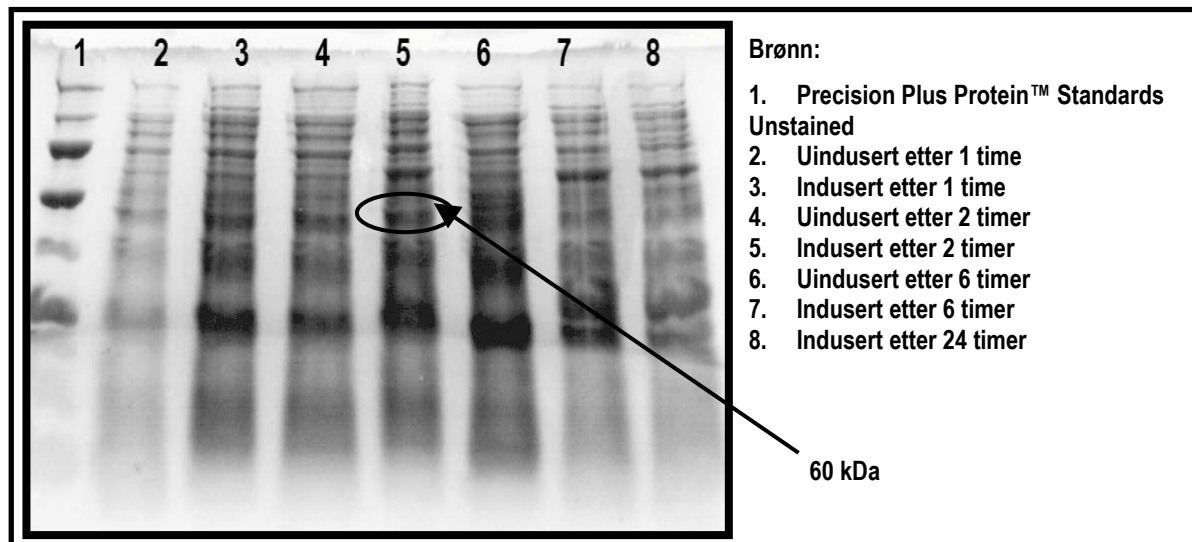
Figur 4.31 viser noe sterkere bånd av proteiner på omkring 20, 30 og 40 kDa. På celleekstrakt fra BL21 celler med vektor uten innskudd, ble disse båndene ikke observert.

Det ble forsøkt å uttrykke *groEL1* ved hjelp av pET 101/ D TOPO vektoren og vertscellen BL21-CodonPlus (DE3) RP. BL21-CodonPlus (DE3) RP vertscellen inneholder ekstra kopier av *argU* og *proL*, genene som koder for tRNA som gjenkjenner kodonene for arginin; AGA og AGG, og kodonet for prolin; CCC. BL21-CodonPlus (DE3) RP vertscellen ble transformert med pET 101/ D TOPO vektoren og dyrket på LB skåler som inneholdt ampicillin og kloramfenikol. Genene som koder for ekstra tRNA ligger på et plasmid som også bærer kloramfenikol-resistens. For å undersøke koloninene for innskudd, ble PCR utført med det vektorspesifikke primersettet T7 og T7 revers. På figur 4. 32 kan resultatet av PCR reaksjonen sees. I brønnene med PCR på kolonier, kan to klare bånd sees. Disse båndene er avmerket på bildet og vil bli omtalt i diskusjonen.



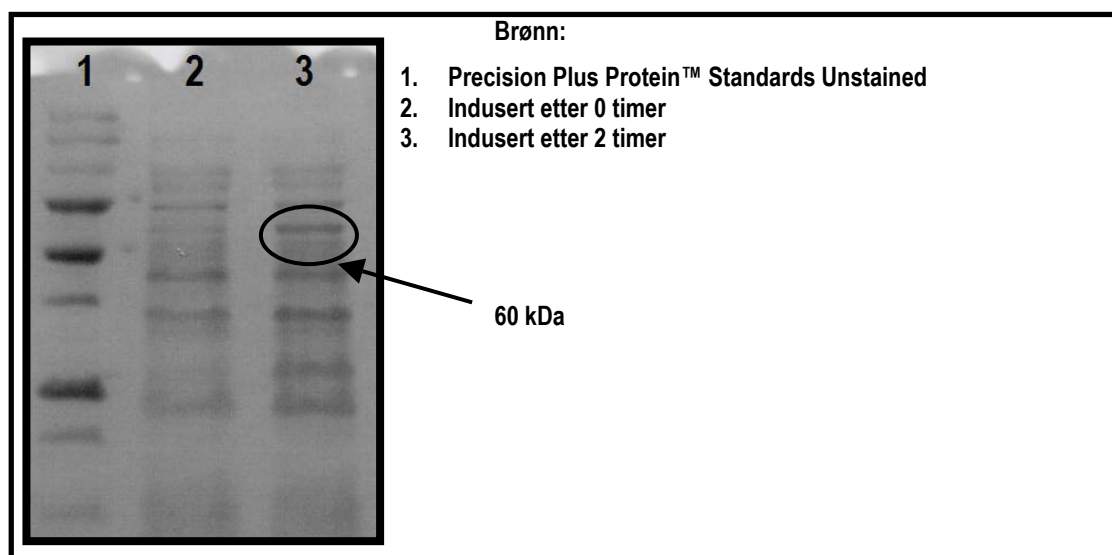
Figur 4.32 PCR på kolonier med *groEL1* innskudd

Kolonier som inneholdt innskudd ble dyrket som overnatt-kulturer og induisert, som beskrevet i protokollen og punkt 3.10. Celleekstraktet fra induerte BL21-CodonPlus (DE3) RP celler ble analysert på SDS-PAGE. Figur 4.33 og 4.34 viser bilder av SDS-PAGE med celleekstraktet fra disse cellene. I brønnene med de induerte prøvene kan et bånd på ca. 60 kDa sees. Båndene er merket av med piler.



Figur 4.33 Bildet viser uttrykk av 60 kDa protein.

BL21-CodonPlus (DE3) RP ble også induisert i kun 2 timer. En nullprøve ble tatt ut før cellene ble induisert med IPTG, og induert prøve tatt etter 2 timer på risteinkubator ved 37 °C. 20 µl av cellekulturen ble tatt ut, tilsatt 10 µl 5x SDS buffer og kokt 5 min ved 95 °C. Prøvene ble deretter analysert på SDS-PAGE. Resultatet er vist på figur 4.34.



4.34 Bildet viser uttrykk av 60 kDa protein.

5 DISKUSJON

I denne oppgaven ble resten av sekvensen til chaperoninet *groESL1* isolert og sekvensert, og det ble vist at genene i operonet, *groES1* og *groEL1*, lar seg uttrykke i pET vektorer med BL21star og BL21 Codon Pluss som vertsceller.

5.1 To operoner for *groESL*

Den termofile grønne glidende bakterien *C. aurantiacus* har to operoner for chaperoninet GroESL, *groESL1* og *groESL2*. *GroESL2* operonet er på 1986 bp. Det er satt sammen av *groES2* på 303 bp og *groEL2* på 1636 bp. Genene er adskilt av en 47 bp lang ikke-kodende sekvens. Sekvensen oppstrøms for *groESL2* har CIRCE sekvensmøtivet og antatte husholdningspromoterer -35 og -10. Promoter -10 sekvensen på TACTAT er noe avvikende fra konsensus sekvensen TATAAT. Shine-Dalgarno sekvenser finnes oppstrøms både for *groES2* og *groEL2*. CIRCE elementet ligger nedstrøms for de antatte -35 og -10 promoterenelementene og omkring 50 bp oppstrøms for startkodonet til *groESL2*.

Operonet *groESL1* er på 2228 bp, der *groES1* er på 312 bp og *groEL1* på 1646 bp. Genene er separert av en sekvens på bare 5 bp. Oppstrøms for *GroESL1* er det ikke funnet noe CIRCE element. Promotersekvensene, -35 og -10, for *groESL1* hos *C. aurantiacus* ligner på de vanlige husholdningssekvensene og sekvensene for sigmafaktorer som benyttes for varmesjokkgenene hos *E. coli*. Ved sekvenssammenstilling i BLAST, Align Two Sequences, viste de to operonene 88 % likhet.

Hvorfor *C. aurantiacus* har to *groESL*-operoner er ennå ikke klart. Det er vist at *groESL1* uttrykkes i økt grad ved bakteriens maksimumstemperatur. Ekspresjon av *groESL2* er enda ikke påvist *in vivo*, men genene i operonet lar seg uttrykke hver for seg, ved hjelp av pET vektorer (Benestad, hovedfagsoppgave UIO, 2005).

Et spørsmål i forbindelse med at organismen har flere kopier av slike gener, er om genene har ulik regulering og/eller funksjon. I *C. aurantiacus* er CIRCE elementet funnet i operonet *groESL2*, men ikke i *groESL1*. Flere forsøk er utført for å finne ut mer om CIRCE elementets rolle i regulering av varmesjokkproteiner. En hypotese er at CIRCE reguleres gjennom en repressor kalt HrcA. HrcA bindes til CIRCE elementet og reduserer transkripsjonen av genet eller operonet nedstrøms for CIRCE elementet. Det er funnet en HrcA homolog i genomet til *C. aurantiacus*. Oppstrøms for HrcA genet, er det dessuten funnet et CIRCE element. Dette kan tyde på at HrcA nedregulerer seg selv under ikke-varmesjokk-forhold.

At de to utgavene av operonet har ulike mekanismer for regulering kan bety at de transkriberes og aktiveres under ulike forhold av stress. Tilstedeværelse av flere varianter av *groESL* gener kan ha betydning ved at de gir bakterien fleksibilitet ved varierende miljøbetingelser og fysiologiske krav. Flere utgaver av de samme genene kan ha flere fordeler slik som økt ekspresjon av det gjeldende proteinet avhengig av forholdene bakterien befinner seg i og dessuten en viss forsikring mot mutasjoner eller andre ødeleggende faktorer. Et av målene med å studere de to genutgavene av *groESL* hos *C. aurantiacus* er for å få kunnskap om disse mekanismene.

5.2 Genbiblioteket

I oppgaven ble det brukt et bibliotek som var konstruert tidligere av Christin Vanberg (hovedoppgave, UIO, 2001). Biblioteket var konstruert fra genomisk DNA fra *C. aurantiacus*, etter at dette var kuttet med *Sau3AI*. Fragmentene ble ligert inn i Lambda ZAP som var behandlet med *Bam*HI, et restriksjonsenzym som gir de samme klebrige endene som *Sau3AI*. Biblioteket ble oppformert ved hjelp av vertscellen *XL1 Blue MRF'*. En ulempe ved oppformering av et bibliotek er at noen fagkloner kan gå tapt da rekombinante fagpartikler har en tendens til å vokse med ulik hastighet. Variasjoner i DNA størrelse og sekvens kan virke inn på replikasjonen til rekombinante fag slik at enkelte kan få økt frekvens mens andre minskes eller mistes helt (Wyman og Werman, 1978). Da det ikke var noen seleksjonsmarkør kan også fag uten innskudd favoriseres. Oppformering av et bibliotek bør derfor ikke foretas mer enn en gang. Biblioteket bør heller ikke såes ut med for høy tetthet da enkelte rekombinanter kan gå tapt på grunn av konkurrerende vekst. Før biblioteket ble tatt i bruk ble det testet gjennom blå hvit bakgrunnsseleksjon for å se på forholdet mellom rekombinante og ikke-rekombinante fag. Det ble også testet i PCR med vektorspesifikke primere som flankerte vektorarmene for å se på størrelsen til innskuddene. Genspesifikke primere spesifikke for *groESL1*, ble testet nøye på bibliotek for å se at biblioteket inneholdt sekvensen av interesse. Alle testforsøkene tydet på at biblioteket kunne brukes til det formålet som var ment i denne oppgaven, å isolere *groESL1*.

Biblioteket som ble brukt var delvis kuttet, for å forhindre at noen fragmenter ble utelatt og for å lage overlappende klon som gjør primer-walking mulig. Gjenkjenningssettet til *Sau3AI* forekommer ofte slik at sannsynligheten for lange sekvenser uten kuttsetter er liten. Forutsetningen for å finne riktig klon i et bibliotek er at fragmentene er store nok til å kunne inneholde hele sekvensen og at hele sekvensen befinner seg på en klon. Fragmenter av størrelsen 3-5 kb var derfor brukt til å konstruere biblioteket. Når størrelsen på fragmenter bestemmes må man ta hensyn til størrelsen på sekvensen man leter etter, hvilken størrelse som lar seg hente opp med metoden man har tenkt til å bruke og hva som lar seg sekvensere ut ifra hvilke midler og metoder man har tilgjengelig. Fragmenter med en størrelse på over 4 kb kan bli vanskelig å amplifisere i PCR og å sekvensere, men vil øke sannsynligheten til å finne sekvensen av interesse på et klon. Det kan derfor være hensiktsmessig å klon mindre fragmenter og heller undersøke flere klon for det ønskede genet. Problemet med denne oppgaven var at hele sekvensen til *groESL1* ikke var kjent. Dersom delen av sekvensen man ikke kjente havnet på en egen klon, ville den være vanskelig å få tak i.

Rekombinante kloner fra bakteriofager som inneholder et ønsket gen kan isoleres ved hjelp av hybridisering og/eller PCR. Hybridisering er en vanlig metode, men et stort problem med denne metoden er å finne en egnet probe. Det er også vanskelig å oppnå hybridiseringsbetingelser som tillater at bare perfekte komplementære hybrider vil dannes. Det er ikke uvanlig at oligoprober med korte sekvenser har en tendens til å hybridisere uspesifikt til høymolekylært DNA (Sambrook et al., 1989). Sannsynligheten øker betraktelig for å kunne isolere et ønsket gen hvis en spesifikk probe er tilgjengelig. I og med at chaperoner er konserverte proteiner hadde det vært mulig å få tak i en probe til *groEL*. GroEL har konserverte områder som har vært brukt som utgangspunkt for konstruksjon av universelle degenererte primere (Rusanganwa et al., 1992). Det ville vært mulig å bruke en spesifikk probe til å isolere det ønskede genet ved å så ut biblioteket, utføre plakk-løft, hybridisere osv., men isolering av gener kan selv med spesifikke prober være vanskelig og tidkrevende. Det er også kostbart på grunn av radioaktive isotoper og membraner. I tillegg har *C.*

aurantiacus 2 utgaver av genet *groESL*, og sannsynligheten for å isolere *groESL2* istedenfor *groESL1* med prober utformet fra konserverte domener hos chaperoniner, var derfor stor. Basert på dette ble det derfor bestemt at *groESL1* skulle isoleres ved hjelp av PCR. I tillegg var *groESL2* isolert ved hjelp av denne metoden i 2001 (Vanberg, hovedfagsoppgave UIO, 2001).

Ved søk etter positive bakteriofagkloner med både metoden til King og metoden til Yu og Bloem ble positive kloner påvist med PCR med genspesifikke primere. Amplifikasjon med de genspesifikke primerene ville gi opphav til bånd på 420 bp og 460 bp på agarosegel, avhengig av hvilket sett som ble brukt. Fragmentet ble sekvensert før det ble brukt i søk i biblioteket, for å forsikre om at riktig sekvens ble amplifisert. Positiv kontroll ble brukt i hvert søk. Det var anbefalt i protokollen til Yu og Bloem å sekvensere PCR produktet for hver positive påvisning. Ved amplifikasjon av positive kloner lå alltid det tilsvarende båndet på linje med den positive kontrollen. Fragmenter av riktig størrelse ble derfor antatt å være den ønskede sekvensen og ble derfor ikke sekvensert for hvert søk.

Metoden til King ble i denne oppgaven bare brukt til å påvise sekundære subbibliotek. Hvis man regner med at en plakk består av omtrent 10^7 bakteriofag, vil det på en skål der det er 100 plakk være $1,0 \times 10^9$ bakteriofag. I metoden til King skulle SM buffer helles over skålene med plakk, slik at plakken diffunderer ut i bufferen. Denne bufferen som inneholdt bakteriofag var subbiblioteket og skulle brukes til å så ut nye skåler. Man kan regne med at ikke alle bakteriofagene diffunderer ut i SM-bufferen. Regner man et svinn på 40 % vil SM-bufferen inneholde ca. $4,0 \times 10^8$ bakteriofag. Når man videre sår ut subbiblioteket på en agarskål med 50 fag er det liten sannsynligheten for at en positiv klon vil være blant de 50 fagene. Som nevnt vil sannsynligvis ikke alle bakteriofagene til en plakk diffundere over i SM-bufferen. Noen plakk vil også være større enn andre, og det vil være flere fag tilstede fra de større plakkene. SM-bufferen vil altså inneholde færre fag enn det som tilsvarer en plakk, og noen kloner vil være overrepresentert i forhold til andre. Noen bakteriofagkloner kan også mistes ved gjentatt amplifikasjon av subbibliotekene, da rekombinante fagpartikler ofte vokser med ulik hastighet. Problemstillingen er lik den som skjer ved amplifikasjon av hele biblioteker. I metoden til King blir subbibliotekene amplifisert på lignende måte for hvert søk. Metoden ble derfor kun brukt til å snevre biblioteket såpass inn at det egnet seg til bruk i metoden til Yu og Bloem.

I metoden til Yu og Bloem ble det brukt en Hybond N membran til å legge over en skål sådd ut fra det primære subbiblioteket funnet med Kings metode. Fagen festet seg til membranen, og membranen ble deretter skyllet i SM-buffer. Dert var viktig at det ikke ble brukt en positivt ladet membran, da slike membraner kan holde for godt på fagene og gjøre det vanskelig å skylle dem av. SM-bufferen med fag, det såkalte fageluatet, ble brukt som templat i PCR.

Metoden til Yu og Bloem er basert på å trinnvis eliminere ikke-positive kloner på en agarskål, uten å overføre fagen til nye skåler. Ved å ikke amplifisere ved hvert søk unngås problemet med at kloner kan gå tapt. I denne oppgaven ble et positivt subbibliotek fra primærsøket fra metoden til King brukt til å så ut skålene. Fordelen med å bruke et subbibliotek fremfor selve biblioteket er at antall fag er mindre, og bakteriene kan derfor såes ut på færre skåler med en lavere tetthet av fag. Det er viktig at tettheten av plakk per skål ikke er for høy, slik at det lett kan plukkes ut en enkelt plakk i siste trinn.

I metoden deles membranen inn i sektorer ved hjelp av en skalpell. Er man uheldig kan man skjære tvers over en plakk. Et positivt signal vil da kunne gis for begge de delte seksjonene, men plakken vil være noe ødelagt og kan være vanskelig å identifisere. Det er derfor hensiktsmessig å undersøke flere positive skåler parallelt. Man kan velge

hvor mange seksjoner man vil dele skålene inn i, men jo flere seksjoner desto større er sannsynligheten for at man skjærer over positive kloner. Man vil allikevel dele opp skålene mange nok ganger til at det ikke blir alt for mange plakk å plukke i siste trinn. I denne oppgaven ble omtrent 20 potensielle plakk plukket, og PCR utført for å identifisere en enkelt positiv klon. I protokollen til Yu og Bloems metode er det beskrevet i siste trinn at den siste membran-sektoren skal hybridiseres. Metoden ble modifisert slik at det ble unødvendig med hybridisering. Plakkene ble heller plukket og løst i SM-buffer, og SM-buffren brukt som eluat i PCR. Dette fungerte svært godt. Ved rette betingelser kan positive klon identifiseres etter et par dager.

Ved hjelp av metoden til King og metoden til Yu og Bloem ble en forlengelse av sekvensen til *groESL1* funnet, isolert og sekvensert. Denne klonen ble kalt Plakk2. Ved nærmere analyse av sekvensen ved hjelp av programmet ExPacy-Transeq ble det imidlertid funnet at dette ikke kunne være hele sekvensen til operonet. Denne sekvensen ga ikke et protein på 60 kDa som er den forventede størrelsen til GroEL proteinene. Sekvensen inneholdt heller ikke stopp-kodonet TAA.

Det ble også gjort forsøk på å isolere hele *groESL1* direkte fra biblioteket ved bruk av kombinasjonen genspesifikk primer med vektorspesifikk primer T7. Resultatet var som forventet, bånd av ulike størrelser avhengig av hvor restriksjonsenzymet hadde kuttet. De fleste av båndene hadde størrelsen 1500 bp og kortere. Klare bånd ble rensset fra gelen og sekvensert. Ingen av dem ga noe resultat videre utover informasjonen fra Plakk2.

Videre ble det prøvd å isolere operonet ved hjelp av "primer-walking" med primere utformet fra sekvensen til Plakk2, der primersettene ble brukt i nye runder med metoden til King og Yu og Bloem. En fortsettelse på *groESL1* sekvensen ble ikke funnet. En sekvenseringsprimer ble utformet etter sekvensen til *groEL1* som allerede var kjent og lå ute på NCBI. Denne primeren var 35 nukleotider lang og ble brukt på genomisk DNA direkte i sekvenseringsreaksjonen. Resultatet var mye bakgrunnsstøy, og et veldig kort kromatogram (kromatogrammet er å finne i Appendiks). Det var ikke mulig å få noe informasjon ut av sekvenseringen. En primer ble utformet fra *groEL2*, Gro2, og ble forsøkt sammen med den *groEL1* spesifikke foroverprimeren for1. Heller ikke dette førte frem til noe resultat. Analysen av PCR reaksjonen på agarosegel viste korte bånd og uspesifikk binding av primere til templat. Selv om *groEL1* og *groEL2* har en identitet på 88% vil primeren binde seg sterkere til det genet den er laget til. Det ble forsøkt å utforme en primer fra konsensussekvensen til chaperoniner fra ulike bakterier. Nukleotidsekvensene ble sammenstillt ved hjelp av programmet GeneDoc. Chaperoniner er meget konserverte og med høy sekvenslikhet, men det var likevel ikke mulig å finne et godt nok konservert område i slutten av konsensussekvensen der en primer kunne utformes.

Fordi genomet til *C. aurantiacus* ikke var ferdig sekvensert da denne oppgaven ble påbegynt, ble det antatt at genomstørrelsen til denne organismen var omtrent 4×10^6 bp, som er genomstørrelsen til *E.coli*. Ut ifra denne størrelsen var genbiblioteket konstruert, og antall plakk nødvendig for å representere hele genomet beregnet. Det viste seg imidlertid da genomet til *C. aurantiacus* var ferdigsekvensert at genomet var på $4,9 \times 10^6$ bp, det vil si noe større enn det man antok da biblioteket ble konstruert. Dette forandret forutsetningene noe. Ut ifra formelen til Clarke og Carbon, som ble brukt til å anslå biblioteket i første omgang, ville sannsynligheten for at et bestemt DNA fragment var tilstede i biblioteket være omkring 0,75 %. Dette visste man ikke da biblioteket ble brukt til å søke etter *groESL1*, og det kan ha hatt betydning for resultatet i og med at hele sekvensen til *groESL1* ikke ble funnet ved hjelp av biblioteket. Metoden til King (noe modifisert) og metoden Yu og Bloem ga gode resultater da de ble brukt til å isolere *groESL2* fra

det samme genbiblioteket (Vanberg, Hovedfagsoppgave UIO, 2001). Hadde det ikke vært for de mange restriksjonssetene i *groEL1* ville den også her antagelig vis ført frem til hele sekvensen.

5.3 Sekvenseringsresultatene

Å finne en positiv klon med hele *groESL1* innskuddet viste seg å være problematisk. Da det viste seg fra sekvenseringsresultatet at alle sekvensene stoppet i samme område, ble sekvensen undersøkt *in cilico*. Det viste seg at området fragmentet stoppet opp i inneholdt flere *Sau3AI* kutteseter. Trolig er det kutteseter her som er mindre metylert, noe som gjør dem mer mottagelige for å bli kuttet. DNA fra *C. aurantiacus* er kjent for at det ikke lett lar seg kutte av visse restriksjonsenzymmer. Det har vært foreslått at bakterien har et *dam*-lignende metyleringssystem (Shiozawa, 1995). Et gen som koder for en metyltransferase har blitt isolert og bekrefter dette (Gruber og Bryant, 1998). Da biblioteket ble konstruert var valg av restriksjonsenzym basert på denne kunnskapen. Programmet Webcutter ble brukt på hele *groEL1* sekvensen for å finne ut hvor i sekvensen *Sau3AI* kuttet. Det viste seg at i området sekvensene stoppet opp var det i alt 17 kutteseter for dette restriksjonsenzymet. Valg av restriksjonsenzym var derfor ikke heldig i forbindelse med denne oppgaven. Selv om det kromosomale DNAet bare ble delvis kuttet, er sannsynligheten for at et av de 17 kuttesetene blir kuttet stor.

Høsten 2004 lå genomet til *C. aurantiacus* ferdigsekvensert på NCBI i form av "shot-gun" sekvenser. Da Plakk2 ble sekvensert og ga en fortsettelse på sekvensen som allerede var kjent, ble denne sekvensen søkt med i BLAST. Det ble lett etter et klon der som ville gi de siste hundre basene som manglet. Sekvensen ble delt opp i mindre fragmenter, 50-100 bp, som ble brukt i nye søk i databasene til NCBI. En klon fra "shot-gun" sekvenseringen, Chlo01_574, viste 100 % likhet med Plakk2. Sekvensen fra Chlo01_574 ble satt sammen med Plakk2 og Chlo_724. Denne sekvensen ble analysert i Expacy og proteinproduktet viste en størrelse på 59,7 kDa. Sekvensen inneholdt også stopp-kodonet TAA. Sekvensen mellom Chlo_724 til Chlo01_574 som kom fra sekvenseringen av Plakk2 ble imidlertid ikke funnet i databasene. Det kan se ut til at denne sekvensen tilhører et av glippene i "shot-gun" sekvenseringen. Dette er sannsynlig i og med at dette området var vanskelig å sekvensere og inneholdt mange kutteseter. Sekvenser som ved agarose gelelektroforese viste lengre sekvenser enn det som kom tilbake fra sekvenseringen kan tyde på vanskelige tredimensjonale strukturer i dette området. Sekvenseringen som var blitt gjort av JGI og som lå ute på NCBI stoppet også her. Den fullstendig sekvensen til *groESL1* er vist i figur 4.24.

5.4 Isolering, transformering og ekspresjon av *groES1* og *groEL1*

Ut fra sekvensene som var funnet ble det utformet primere til å isolere *groES1* og *groEL1*. Primersettene ble brukt i PCR med både genomisk DNA og bibliotek som templat. Primersettet til *groES1* ga gode resultater på begge templatene, og da PCR produktene ble sendt til sekvensering, viste de å være korrekte. Primersettet til *groEL1* ga bare resultat da genomisk DNA ble brukt som templat. PCR produktet viste seg å være av ønsket størrelse og da produktet ble sekvensert, stemte det med både 5'- og 3'- ende av den sekvensen man antok var sekvensen til *groEL1*. At primersettet til isolering av *groEL1* ikke ga noe resultat i form av isolerte bånd da det ble brukt på bibliotek som templat styrket teorien om at dette genet i sin helhet ikke befant seg på en klon.

De to genene, *groES1* og *groEL1* skulle uttrykkes i ekspresjonsvektorer. Genene *groES2* og *groEL2* fra *C. aurantiacus* hadde med hell blitt uttrykt henholdsvis ved hjelp av ekspresjonsvektorene pET 101/D TOPO og pET 102/D TOPO, og med BL21 Star som vertsceller (Benestad, Hovedoppgave UIO, 2005)

Vektoren pET 102/D TOPO er en vektor der det heterologe genet blir uttrykt som et fusjonsprotein sammen med proteinet tioredoksin. Denne vektoren ble brukt da *groES2* er på bare 10 kDa og er vanskelig å synliggjøre på vanlig SDS_PAGE. Det ble oppnådd ekspresjon av *groES1* som fusjonsprotein, på 24,5 kDa, da dette genet ble satt inn i vektoren pET 102/D TOPO og uttrykt i BL21 Star celler. Produktet til *groEL1* er et protein på 60 kDa. Det ble først forsøkt å uttrykke *groEL1* som et fusjonsprotein i pET 102 D TOPO vektoren. Det ble ikke oppnådd noen ekspresjon bortsett fra fusjonsproteinet tioredoksin. Plasmidet med innskuddet ble sekvensert og viste at innskuddet var orientert riktig vei i vektoren samt var i korrekt leserramme. At tioredoksin ble uttrykt tydet på at induksjonen av cellene hadde virket. Båndet nederst på bildet av gelen (figur 4.35) viser ekspresjon av tioredoksin, 13 kDa. Hadde vektoren ikke inneholdt innskudd, ville produktet vært noe større da transkripsjonen ville inneholdt V5-epitop og histidinhale. Ekspresjonsproduktet av *groEL1* med fusjonsproteinet ville hatt en størrelse på 73 kDa.

Når ekspresjon med pET vektorene ikke fungerer, skyldes dette i 5 % av tilfellene at produktet er toksisk for vertscellen (Invitrogen). Ved å tilsette glukose i mediet under ekspresjonsforsøket vil ekspresjonen av T7 RNA-polymerasen undertrykkes, og mindre av det heterologe genet dermed uttrykkes. Genprodukter som påvirker vertscellens vekst ved lave konsentrasjoner, regnes som toksiske. Toksiske produkter kan føre til tap av plasmid og/eller langsom vekst i forhold til kontrollen. Det ble ikke observert noen langsommere vekst hos IPTG-induserte celler som inneholdt vektor med innskudd. Det ble likevel forsøkt med 1 % glukose i mediet, men det hadde ingen innvirkning på ekspresjonen. Det ble derfor antatt at *groEL1* ikke var toksisk for vertscellen. I tillegg hadde *groEL2* blitt uttrykt i det samme systemet og viste ingen tegn til å være toksisk (Benestad, Hovedoppgave UIO, 2005).

Det ble vurdert om *groEL1* fusjonert med tioredoksin ble for langt og inneholdt for mange sekundære strukturer som hindret RNA polymerasen. Dersom genet som settes inn i vektoren inneholder sekundære strukturer kan det være vanskelig å få noen ekspresjon. Det ble derfor forsøkt å uttrykke *groEL1* uten fusjonsprotein med vektoren pET 101/D TOPO. Ved å øke tiden for induksjon fra 6 til 12 timer og senke temperaturen fra 37°C til 25 °C, kan det hjelpe RNA-polymerasen å lese gjennom vanskelige deler av genet istedenfor å "hoppe over" kompliserte strukturer. Dette ble forsøkt, men et produkt på rundt 60 kDa ble ikke observert. Av bildene som ble tatt av gelene (figur 4.37) kunne derimot proteiner på størrelsene 20, 30 og 40 kDa observeres. Disse hadde en økt ekspresjon i de IPTG-induserte cellene.

Ekspresjon av heterologe gener i *E. coli* kan være et problem hvis kodonbruken i det rekombinante genet er ulik den i vertscellen. Høy grad av ekspresjon av gener med kodoner som sjeldent brukes i *E. coli* kan skape problemer under ekspresjon fordi det ikke er nok tRNA tilstede, dette kalles kodon-bias. Dette kan medføre at translasjonen av mRNA blir forsinket, proteinene raskere degradert, rammebytte og/eller kodon-substitusjon og feilinkorporering som ødelegger den funksjonelle karakteren til proteinet. Det kan også oppstå problemer med for tidlig terminering av transkripsjon og/eller translasjon, noe som medfører proteinprodukter av ulike lengder. Dette problemet er best dokumentert for noen av kodonene for arginin, AGA og AGG, som er de to minst brukte kodonene i *E. coli*. Noen kodoner for isoleucin (AUA), leucin (CUA) og prolin (CCC) er også kjent for å ha en innvirkning på mengden og

kvaliteten av protein produsert i *E. coli* vertsceller. For å undersøke om *groEL1* inneholdt kodoner som sjeldent brukes i *E. coli*, ble programmet Rare Codon Calculator brukt. I *E. coli* er det særlig noen av kodonene for arginin, glycin, isoleucin og prolin som har få tilhørende tRNA-molekyler. Det ble funnet 4 sjeldne kodoner for aminosyren prolin (CCC), og 10 sjeldne kodoner for aminosyren arginin (AGG, AGA, CGA). Dette er kodoner som vanligst skaper problemer i GC rike genomer. Av disse befant 8 av dem seg i de 400 siste basene i genet (figur 4.36).

I *groEL1* lå de aller fleste sjeldne kodonene i den C-terminale enden av den kodende sekvensen. Ekspresjon av det heterologe genet blir spesielt påvirket når de sjeldne kodonene er i den N-terminale enden av den kodende sekvensen eller hvis de er samlet på et område av sekvensen. Genet ble delt opp i 4, etter hvor de sjeldne kodonene lå i sekvensen. Delene ble brukt i programmet ExPacy-Transeq for å se på hvordan produkt sekvensen ville danne dersom translasjonen stoppet for tidlig på grunn av sjeldne kodoner. Resultatet viste proteinprodukter på 20, 30 og 40 kDa. Celleekstraktet fra induerte BL21 celler ble analysert på SDS-PAGE. På disse gelene var produkter av nettopp disse størrelsene å observere (se figur 4.37).

Fordi de sjeldne kodonene var spredt over et så stort område på genet var det vanskelig å utforme primere som kunne isolere genet og innføre punktmutasjoner for å forandre de sjeldne kodonene til mer vanlige kodoner, ved hjelp av PCR. I et slikt tilfelle kunne for eksempel A blitt erstattet med U slik at det sjeldne CGA-kodonet for arginin ble gjort om til CGU, som er et mer vanlig kodon i *E. coli*. En annen måte å løse kodon-bias problemet er å benytte seg av vertsceller som inneholder ekstra kopier for tRNA genene det er mangel på. *E. coli* stammen BL21-CodonPlus® fra Stratagene er en slik vertscelle. BL21-CodonPlus® stammene har ekstra kopier av sjeldne *E. coli* tRNA-genene, og tillater derfor en høyere ekspresjon av gener med uvanlige kodon enn andre *E. coli* verstceller. De ekstra tRNA genene sitter på et plasmid som i tillegg inneholder kloramfenikol-resistens. BL21 CodonPlus (DE3)-RP er for GC rike gener og har derfor ekstra kopier av tRNA for prolin og arginin. Det lyktes å uttrykke *groEL1* med pET 101/D TOPO vektoren i BL21-CodonPlus-(DE3)-RP cellene.

5.5 Sekundære strukturer i *groEL1*

Sekundære strukturer kan gjøre sekvensering av nukleotidsekvenser vanskelig. Det kan føre til ingen signaler, svake signaler, forkortede sekvenser eller bakgrunnsstøy, som en følge av at sekvensene kommer ut av kapillærrørene i sekvenseringsmaskinen samtidig fordi sekvensene har bundet seg sammen. Sekundære strukturer kan dannes under PCR reaksjonen der DNAet gjøres enkelttrådig til sekvensering, eller under selve sekvenseringen som foregår over 4-6 timer ved 44 °C. Som beskrevet i innledning og metoder (punkt 1.5 og 4.6) kan sekundære strukturer i enkelttrådig DNA analyseres på agarosegel med etidiumbromid. Fordi etidiumbromid interkalierer med basene i dobbeltrådig DNA vil mengde DNA som synliggjøres på gelen under uv-lys kunne være et mål på mengden av dobbeltrådig DNA vil dannet fra det enkelttrådig DNAet. Strukturene i *groEL1* ble analysert på denne måten (punkt 3.12). Det viste seg på bildene av agarosegelene (figur 4.28) at programmet der temperaturene var høyest, program 3, hadde færrest synlige bånd. Sekvenseringsresultatene fra denne prøven ga også det lengste og beste kromatogrammet. Ved bruk av standard oppsett for sekvensering av PCR produkter ble det nesten ingen synlige bånd på bildet av agarosegelen, figur 4.29, men disse prøvene ga de dårligste sekvenseringsresultatene. Resultatet kan tyde på at det med program 3 ble dannet lange enkelttrådig DNA sekvenser, som dannet noen sekundære strukturer, mens det i standardoppsettet ble

korte sekvenser uten sekundære strukturer. Dette kan stemme fordi det var i siste del av *groESL1* sekvenseringsproblemene oppstod. Polymerasen kan ha problemer med å lese gjennom vanskelige strukturer og glippe slik at resultatet blir forkortede PCR produkter. Ved sekvensering av PCR produktene fra de 3 forskjellige programmene ble sekvenseringen utført ved 50 °C. Dette kan ha hindret dannelsen av sekundære strukturer fordi høyere temperaturer bryter bindingene mellom DNA trådene. Det er ikke uvanlig at amplifikasjon av templatet over 1 kb kan gi problemer. DNA kan være GC-rikt og/eller ha sekundære strukturer. Sekundærstrukturene kan være lokalisert i spesielle områder av DNAet. Disse strukturene kan fortsatt være tilstede selv om DNA-stykket utgjør et innskudd i en vektor. Hvis templatet er tvunnet i en løkkestruktur vil polymerasen kunne krysse over strukturen under PCR (Sharma et al., 1992). Amplifikasjon vil dermed gi et produkt som er kortere enn det originale templatet. Korte produkter vil ofte favoriseres i PCR, og hvis noen få slike produkter først dannes, vil disse fort dominere reaksjonen. Nukleotidsekvensen til *groESL1* har et G/C-innhold på 58 %. Der problemene med sekvenseringen oppstod er G/C innholdet på 60 %. Sekvensen ble også analysert i programmet mFold. Resultatet av disse analysene viste at dinukleotidhyppigheten av GC var relativt mye høyere i slutten av sekvensen. GC rike templatet kan skape problemer i forbindelse med PCR da denaturering av templatet kan være vanskelig. Problemer med PCR på templatet som er GC rike og/eller har sekundære strukturer kan ofte løses ved at man tilsetter en forbindelse i PCR reaksjonsblandingen som vil hjelpe med å åpne/denaturere templatet og rette ut strukturene. En slik forbindelse kan for eksempel være DMSO eller glyserol.

Da vektorer som inneholdt *groEL1* ble analysert ved hjelp av PCR, viste resultatet et bånd av korrekt størrelse, 1600 bp, og et bånd 1100 bp. Disse båndene kunne vært renset fra gelen og sekvensert for å finne ut mer om de sekundære strukturene til *groEL1*. Dette ble ikke gjort.

5.6 Forslag til videre arbeid

De to operonene *groESL1* og *groESL2* fra *C. aurantiacus* er sekvensert og isolert. Begge operonene har blitt uttrykt i pET baserte ekspresjonssystemer med *E.coli* BL21 Star og BL21 Codon Pluss som vertsceller. Spørsmålet om hvorfor *C. aurantiacus* har to utgaver av dette genet gjenstår å besvare.

Flere forsøk har blitt gjort for å undersøke dette. Ved varmemstress har det blitt vist økt ekspresjon av *groESL1* (Tooming, Hovedfagsoppgave UiO, 1998). Hvilke forhold som induserer *groESL2* er fortsatt ikke klart. Varmestress, lysstress, oksidativt stress, saltstress, etanolstress samt lave og høye pH verdier har blitt prøvd (Benestad, Hovedfagsoppgave UiO, 2005), men uten resultat. Ved å se nærmere på sekvensene til de to chaperoninene, og genomet til *C. aurantiacus* som nå er ferdigsekvensert, kan man finne ut mer om de to genes reguleringssystemer.

Når et gen oppreguleres øker mengden mRNA for dette genet i cellen. Ved hjelp av metoden Real-Time RT-PCR kan man kvantitativt måle mRNA mengden til spesifikke gener. I denne metoden kopieres mRNAet til cDNA ved hjelp av revers transkriptase. Deretter analyseres cDNAet i Real-Time PCR der det brukes fluorescerende farge som lyser når det bindes til dobbelttrådet DNA. Fluorescensen måles etter hver PCR syklus og blir et mål på mengde mRNA i den originale prøven. På denne måten kan man undersøke hvilke gener som uttrykkes, om de uttrykkes og i hvilken grad. Ved å behandle cellene før RNA ekstraksjonen, kan man undersøke mRNA nivået i cellen under ulike forhold, f. eks. stress. Ved hjelp av denne metoden kan man sammenligne de to *groESL* operonene i *C. aurantiacus* og finne ut mer om reguleringssystemene. Det er funnet sekvenser i genomet til *C. aurantiacus* som viser likhet med *hrcA*, *grpE*,

dnaK og *dnaJ*. Disse genene og deres betydning i varmesjokkresponsen er beskrevet i innledningen (punkt 1.2.2). Oppstrøms for *hrcA* er det også funnet et CIRCE element (Patje, semesteroppgave UiO, 2004).

Det ble i løpet av arbeidet med denne oppgaven undersøkt om det gikk an å sette *groESL1* i sin helhet inn i en ekspresjonsvektor, såkalt koekspresjon. Flere firmaer ble kontaktet, men en slik vektor ble ikke funnet. Problemet med å sette inn et operon er at ekspresjonsvektorene vanligvis bruker vertscellens promotersystem. Når *groES* er transkribert vil polymerasen hoppe av. Fordi det ikke er et promoterområde foran *groEL* vil dette genet ikke transkribes. I *C. aurantiacus* transkribes disse to genene i ett. De to genene i operonet translateres så som to proteiner i cellen. Det som kan være en mulighet er å transformere vertsceller med to plasmider, et som inneholder *groEL* og et som inneholder *groES*. Når de to heterologe genene i plasmidene transkribes og translateres vil chaperoninet GroESL muligens bli funksjonelt. Ved å analysere celleekstrakt fra transformerte og induserte vertsceller på SDS-PAGE kan man finne ut om cellene har tatt opp begge plasmidene og om begge de to heterologe genene har blitt uttrykt. Videre vil det være interessant å dyrke disse cellene ved økende temperatur for å finne ut om det er mulig å gjøre *E. coli* mer varmetabil. Slike forsøk har vært utført der man har introdusert klonet *groEL* i mutanter av *E. coli* som ikke er varmetolerante. Klonet *groEL1* fra *Synechococcus vulcanus* har blitt transformert i *E. coli* og har ført til økt varmetoleranse (Tanaka *et al.*, 1997). Lignende forsøk kan utføres med chaperoninene til *C. aurantiacus*.

REFERANSELISTE

- Altschul, S.F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J. Basic Local Alignment Search Tool. J. Mol. Biol., 215(3) : 403-10.
- Babst, M., Hennecke, H., Ficsher, H-M. 1996. Two different mechanisms are involved in the heat shock regulation of chaperonin gene expression in *Bradyrhizobium japonicum*. Mol Microbiol, 19 : 827-839.
- Bauld, J., Brock, T.D. 1973. Ecological studies of Chloroflexis, a gliding photosynthetic bacterium. Arch. Microbiol., 92:267-284.
- Baird, P.N., Hall, L.M.C., Coates, A.R.M. 1989. Cloning and sequence analysis of the 10 kDa antigen gene of *Mycobacterium tuberculosis*. J Gen. Microbiol., 135 : 931-939.
- Baker, D.J., Gore, M.G. 1998. A chaperonine apical domain from the thermophilic bacterium *Thermus aquaticus*. Biochem. Soc. Trans., 26(3) : S250.
- Baldini, R.L., Avedissian, M., Gomes, S.L. 1998. The CIRCE element and its putative repressor control cell cycle expression of the *Caulobacter crescentus* groESL operon. J. Bacteriol., 180:1632-1641.
- Benestad, I. 2005. Isolering, kloning og ekspresjon av groESL2-operonet fra den termofile bakterien *Chloroflexus aurantiacus*. Hovedfagsoppgave i generell mikrobiologi ved Avdeling for Molekylær biovitenskap, Biologisk Institutt, Universitetet i Oslo.
- Bindon, C., Martindale, J., Mitchell, C. 1998. Biologically-generated primer for PCR: PCR primer of unknown sequence. Nucleic Acids Research Pages 3305-3308.
- Bollag, D. M., Rozycki, M. D., Edelstein, S. J. 1996. Protein methods. 2nd Edition. Wiley-Liss Inc.
- Braig, K., Otwinowski, Z., Hedge, R., Boisvert, D.C., Joachimiak, A., Horwich, A.L., Sigler, P.B. 1994. The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. Nature, 371:578-586.
- Burston, S.G., Ranson, N.A., Clarke, A.R. 1995. The origin and consequences of asymetry in the chaperonin reaction cycle. J. Mol. Biol., 249:138-152
- Burston, S.G., Weissman, J.S., Farr, G.W., Fenton, W.A., Horwich, A.L. 1996. Release of both native and non-native proteins from a cis-only GroEL ternary complex. Nature , 383:96-99.
- Carstens, C. P., Waesche, A. 1999. Codon Bias-Adjusted BL21 Derivatives for Protein Expression Strategies Newsletters (Stratagene) May 1999; vol. 12 #2 pg. 49-51
- Castenholz, R. W. Pierson, B. K. 1995. Ecology of the thermophilic anoxygenic phototrophs. In Blankenship, R.E., Madigan , M. T., Bauer, C. E (ed.). Anoxygenic photosynthetic bacterie. Kluwer Academic Publishers, Dordrecht, the Netherlands
- Chen, S., Roseman, A.M., Hunter, A., Wood, S.P., Burston, S.G., Ranson, N., Clarke, A.R., Saibil, H.R. 1994. Location of a folding protein and shape changes in GroEL-GroES complexes imaged by cryo-electron microscopy. Nature, 371:261-264.

- Danson, M.J., Hough, D.W. 1998. Structure, function and stability of enzymes from the archaea. Trends in microbiology, 6:307-314.
- Ditzel, L., Löve, J., Stock, D., Stetter, K.O., Huber, R., Steinbacher, S. 1998. Crystal structure of the thermosome, the archaeal chaperonin and homolog of CCT. Cell, 93:125-138.
- Dirks, R. M., Pierce, N. A. 2003. A Partition Function Algorithm for Nucleic Acid Secondary Structure Including Pseudoknots. J Comput Chem 24:1664-1677.
- Dirks, R. M., Pierce, N. A. 2004. An Algorithm for Computing Nucleic Acid Base-Pairing Probabilities Including Pseudoknots. J Comput Chem 25:1259-1304.
- Dorfman, D.M. 1993. Amplification of bacteriophage library inserts using polymerase chain reaction. Methods in Enzymology, 218:336-340.
- Ellis, R.J, Hartl, F. U. 1996. Protein folding in the cell: competing models of chaperonin function. FASEB J., 10: 20-26.
- Fayet, O., Ziegelhoffer, T., Georgopoulos, C. 1989. The *groES* and *groEL* heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. J. Bacteriol., 171(3):1379-85.
- Feldman, D.E., Frydman, J. 2000. Protein folding *in vivo*: the importance of molecular chaperones. Curr. Opin. Struct. Biol., 10(1):26-33.
- Fenton, W.A., Kashi, Y., Furtak, K., Horwich, A.L. 1994. Residues in chaperonin GroEL required for polypeptide binding and release. Nature, 371:614-619.
- Frackman, S., Kobs, G., Simpson, D. and Storts, D. 1998. Betaine and DMSO: enhancing agents for PCR. Promega Notes 65: 27.
- Fraser, D.C., Craigmiller, S., Russel, G.C. 1994. Accelerated screening of cDNA libraries using the polymerase chain reaction and southern blotting. Genet. Anal. Tech. Appl., 11(4):87-9.
- Garner, J., Multhaup, G., Tomoyasu, T., McCarty, J.S., Rüdiger, S., Schönfeld, H.J., Schirra, C., Bujard, H., Bukau, B. 1996. A cycle of binding and release of the DnaK, DnaJ and GrpE chaperones regulates activity of the *Escherichia coli* heat shock transcription factor σ 32. EMBO J., 15:607-617.
- Gething, J.L., Sambrook, J.F. 1992. Protein folding in the cell. Nature, 355:33-45.
- Gor, D., Mayfield, J.E., 1992. Cloning and nucleotide sequence of the *Brucella abortus groE* operon. Biochem. Biophys. Acta, 1130:120-122.
- Gruber, T.M., Bryant, D.A. 1998. Characterisation of the group 1 and group 2 sigma factors of the green sulfur bacterium *Chlorobium tepidum* and the green non-sulfur bacterium *Chloroflexus aurantiacus*. Arch. Microbiol., 170:285-296.
- Gupta, R.S., Mukhtar, T., Singh, B. 1999. Evolutionary relationship among prokaryotes (*Heliobacterium chlorum*, *Chloroflexus aurantiacus*, cyanobacteria, *Chlorobium tepidum* and proteobacteria) : implications regarding the origin of photosynthesis. Mol. Microbiol., 32(5):893-906.
- Hartl, F.U. 1996. Molecular chaperons in cellular protein folding. Nature, 381:571-580.

- Hendrick, J.P., Hartl, F.U. 1993. Molecular chaperons functions of heat shock proteins. *Annu. Rev. Biochem.*, 62:349-84.
- Herman, C., d'Ari, R. 1998. Proteolysis and chaperones: the destruction/reconstruction dilemma. *Curr. Opin. Microbiol.*, 1:204-209.
- Holo, H. 1989. *Chloroflexus aurantiacus* secretes 3-hydroxypropionat, a possible intermediat in the assimilation of CO₂ and acetate. *Arch. Microbiol.*, 151:252-256.
- Holo, H., Sirevåg, R. 1986. Autotrophic growth and CO₂ fixation of *Chloroflexus aurantiacus*. *Arch. Microbiol.* 145:173-180.
- Horoviz, A. 1998. Structural aspects of GroEL function. *Current Opinion in Structural Biology*, 8:93-100.
- Horowich, A.L., Weber-Ban, E.U., Finley, D. 1999. Chhaperone rings in protein folding and degradation. *Proc. Natl. Acad. Sci. USA*, 96(29):11033-40.
- Hunt, J.F., Weaver, A.J., Landry, S., Gierasch, L. Deisenhofer, J. 1996. The crystal structure of the GroES co-chaperonin at 2.8 Å resolution. *Nature*, 379:37-45.
- Israel, D.I., 1993. A PCR-based method for high stringency screening of DNA libraries. *Nucleic Acid Research*, 21(11):2627-2631.
- Itzhaki, L.S., Otzen, D.E., Fersht, A.R. 1995. Nature and consequences of GroEL-protein interactions. *Biochemistry*, 34(44):14581-7.
- Jackson, S.E. 1998. How do small single-domain proteins fold? *Fold.*, 3(4):R81-91
- Kad, N.M., Ranson, N.A., Cliff, M.J., Clarke, A.R. 1998. Asymmetry, commitment and inhibition in the GroE ATPase cycle impose alternating functions on the two GroEL rings. *J. Mol. Biol.*, 278(1):267-78.
- King, M.W. 1997. Rapid and nonradioactive screening of recombinant libraries by PCR. *Methods in Molecular Biology*, 67:331-338.
- Klumpp, M., Baumeister, W., Essen, L-O. 1997. Structure of the substrate binding domain of the thermosome, an archaeal group 2 chaperonin. *Cell*, 91:262-270.
- Landry, S.J., Zeilstra-Ryalls, J., Fayet, O., Georgopoulos, C., Gierasch, L.M. 1993. Characterization of a functionally importante mobile domain of GroES. *Nature*, 364:255-258.
- Lardelli, M. 1997. Generation and PCR screening of bacteriophage Lambda sublibraries enriched for rare clones ("the sublibrary method"). *Methods in Mol. Biol.*, 67:345-355.
- Lin, S., Schwarz, F.P., Eisenstein, E. 1995. The hydrophobic nature of GroEL-substrate binding. *J. Biol. Chem.*, 270:1011-1014.
- Løken, Ø., Sirevåg, R. 1982. Evidence for the presence of the glyoxylate cycle in *Chloroflexus*. *Arch. Microbiol.*, 132:276-279.
- Macario, A.J., Lange, M., Ahring, B.K., De Macario, E.C. 1999. Stress genes and proteins in the archaea. *Microbiol. Mol. Biol. Rev.*, 63(4):923-67.

Madigan, M.T., Brock, T.D. 1975. Photosynthetic sulfide oxidation by *Chloroflexus aurantiacus*, a filamentous photosynthetic gliding bacterium, J. Bacteriol., 122:782-784.

Madigan, M.T., Martinko, J.M., Parker, J. 1999. Brock Biology of Microorganisms. 9nd. Prentic Hall, Inc.

Madigan, M.T., Petersen, S.R., Brock, T.D. 1974. Nutritional studies of *Chloroflexus*, a filamentous photosynthetic gliding bacterium. Arch. Microbiol., 100:97-103.

Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A.L., Hartl, F-U. 1991. Chaperonine-mediated protein folding at the surface of GroEL through a molten globule-like intermediate. Nature, 352:36-42.

McBride MJ. 2001. Bacterial gliding motility: multiple mechanisms for cell movement over surfaces. Annu. Rev. Microbiol. 55:49-75.

Mikulik, K., Benada, O. 1993. GroEL-like protein complex of thermophilic bacterium *Thermus aquaticus*. Biochem. Biophys. Res. Commun., 197(2):716-21.

Minder, A.C., Fischer, H.M., Hennecke, H., Narberhaus, F. 2000. Role of HrcA and CIRCE in the heat shock regulatory network of *Bradyrhizobium japonicum*. Bacteriol., 182:14-22.

Morimoto, R.I., Tissieres, A., Georgopoulos, C. 1994. The biology of heat shock proteins and molecular chaperones. (Cold Spring Harbour, N.Y. : Cold Spring Harbour Laboratory Press).

Mullis, K.B. 1990. The unusual origin of the polymerase chain reaction. Sci. Am., 262:56-65.

Munroe, D.J., Loebbert, R., Bric, E., Whitton, T., Prawitt, D., Vu, D., Buckler, A., Winterpacht, A., Zabel, B., Housman, D.e. 1995. Systematic screening of an arrayed cDNA library by PCR. Proc. Natl. Acad. Sci. USA, 92(6):2209-13.

Nakahigashi, K., Ron, E.Z., Yanagi, H., Yura, T. 199. Differential and independent roles of a σ^{32} homolog (RpoH) and a HrcA repressor in the heat shock response of *Agrobacterium tumefaciens*. J. Bacteriol., 181:7509-7515.

Narberhaus, F., Bahl, H. 1992. Cloning, sequencing and molecular analysis of the *groESL* operon of *Clostridium acetobutylicum*. J. Bacteriol., 174:3282-3289.

Patje, W. 2004. The chaperonin regulation in *Chloroflexus aurantiacus*. Semesteroppgave i generell mikrobiologi ved Institutt for Molekylær Biovitenskap, Universitetet i Oslo.

Phipps, B.M., Typke, D., Hegerl, R., Volker, S., Hoffman, A., Stetter, K.O., Baumeiser, W. 1993. Structure of molecular chaperone from a thermophilic archaeobacterium. Nature, 361:475-477.

Pierson, B.K., Castenholz, R.W. 1974. A phototrophic gliding filamentous bacterium of hot springs, *Chloroflexus aurantiacus*, gen and sp. Nov. Arch. Microbiol., 100:5-24.

Ranson, N.A., Dunste, N.J., Burston, S.G., Clarke, A.R. 1995. Chaperonins can catalyze the reversal of the early aggregation steps when a protein misfolds. J. Mol. Biol., 150:581-586.

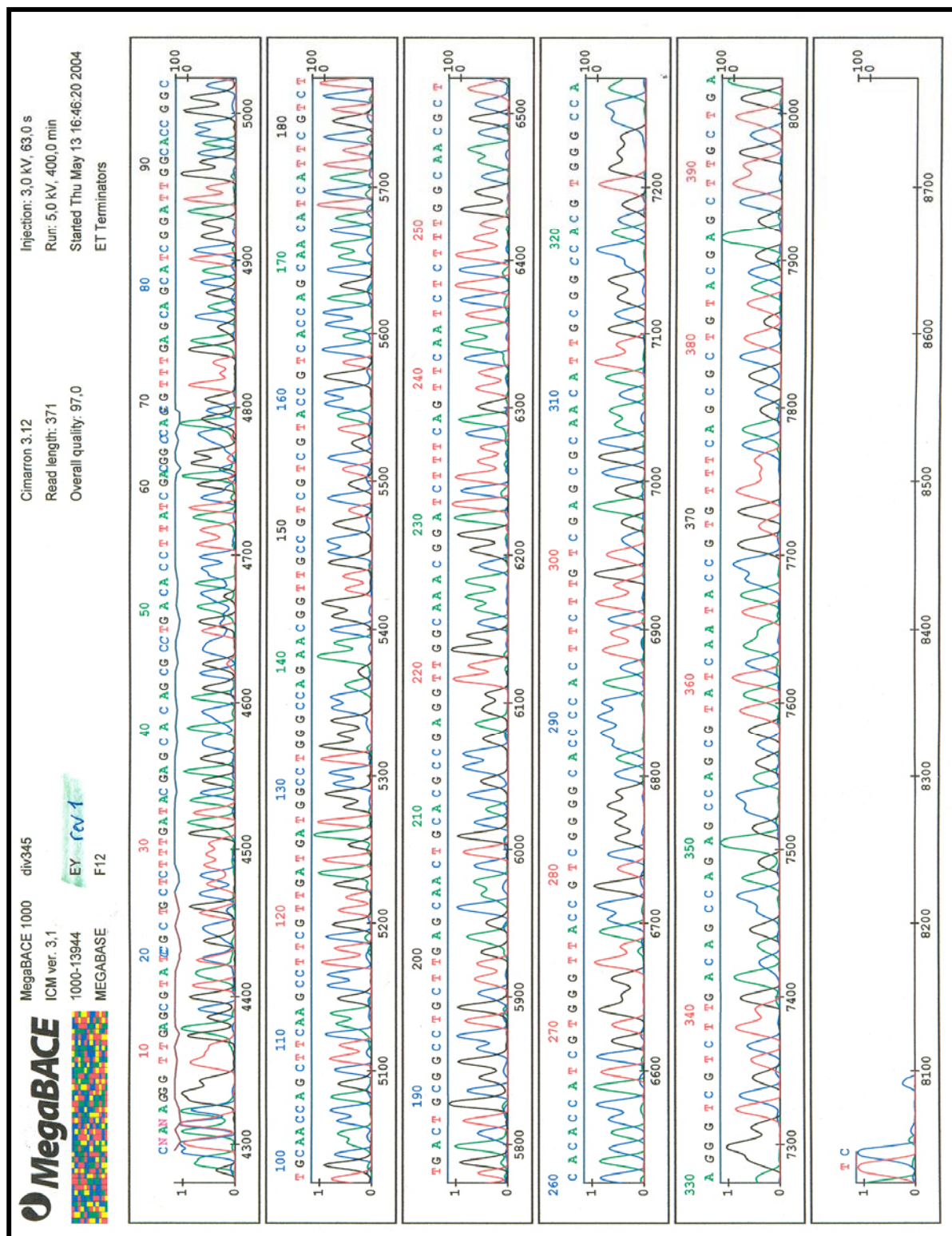
Ranson, N.A., White, H.E., Saibil, H.R. 1998. Chaperonins. Biochem. J., 333:233-242.

Roseman, A.M., Chen, S., White, H., Braig, K., Saibil, H.R. 1996. The chaperonin ATPase cycle: mechanism of allosteric switching and movements of substrate-binding domains in GroEL. Cell, 87:241-251.

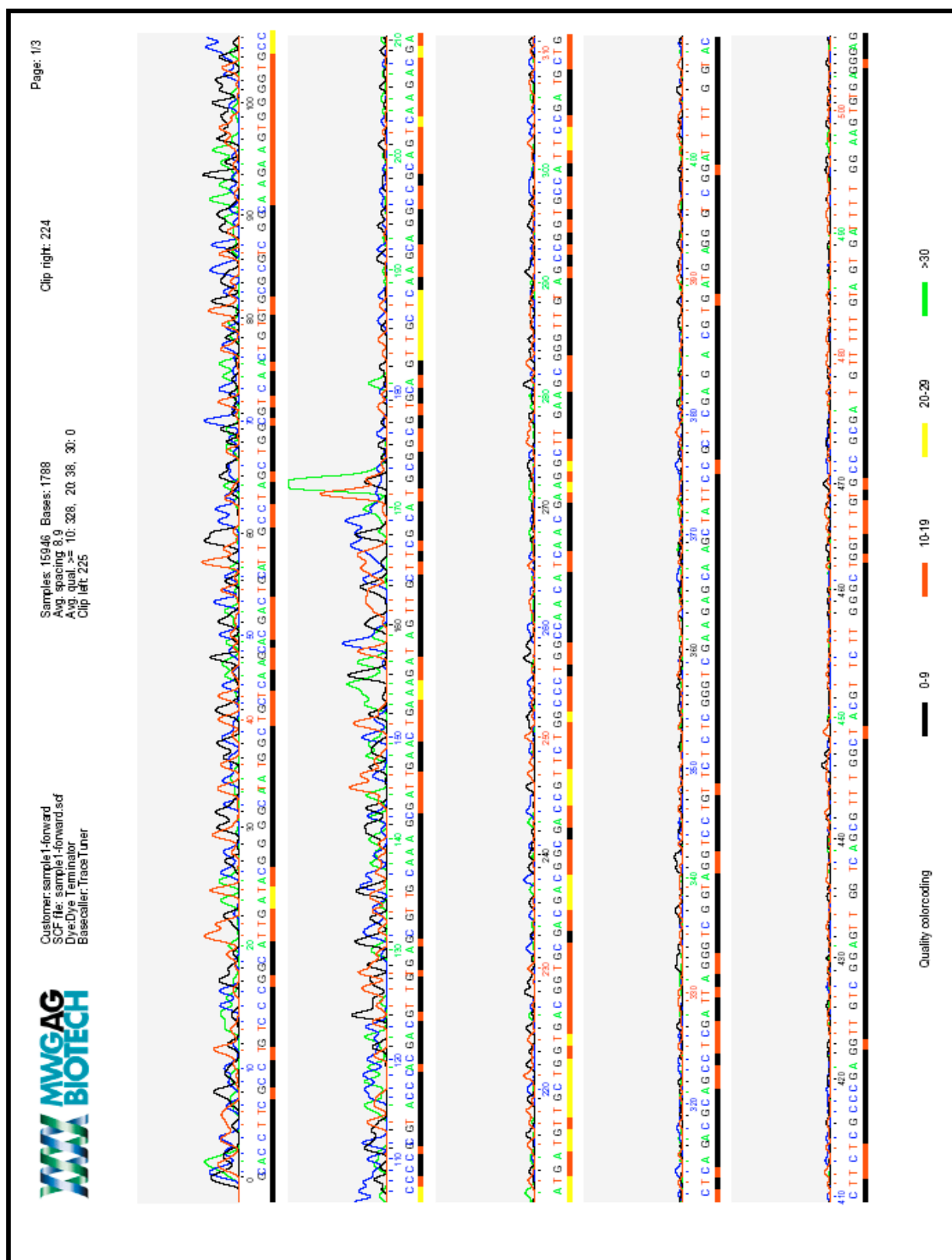
- Roseman, A.M., Bukau, B. 1998. Molecular Chaperones and Folding Catalysts. *Rev. Biochem. J.* V:333
- Rusanganwa, E., Gupta, R.S. 1993. Cloning and characterisation of multiple *groEL* chaperonin-encoding genes in *Rhizobium meliloti*. *Gene*, 126:67-75.
- Rye, H.S., Burston, S.G., Fenton, W.A., Beechem, J.M., Xu, Z., Sigler, P.B., Horwich, A.L. 1997. Distinct actions of *cis* and *trans* ATP within the double ring of the chaperonin GroEL. *Nature*, 388:792-798.
- Sambrook, J., Fritsch, E.F., Maniatis, T. 1989. Molecular cloning: A laboratory manual. 2nd Cold Spring Harbor Laboratory, Cold Spring Harbor NY.
- Sang, J., Thompson, N.L. 1994. An efficient procedure for obtaining long cDNA clones from phage library screening. *Biotechniques*, 17(3):446, 448, 450-1.
- Schon, U., Schuman, W. 1993. Molecular cloning, sequencing and transcriptional analysis of the *groESL* operon from *Bacillus stearothermophilus*. *J. Bacteriol.*, 175:2465-2469.
- Sharma, J.K., Gopalkrishna, V., Das, B.C. 1992. A simple method for elimination of unspecific amplifications in polymerase chain reaction. *Nucleic Acids Res.*, 20(22):6117-8.
- Short, J.M., Fernandez, J.M., Sorge, J.A., Huse, W.D. 1988. Lambda ZAP: a bacteriophage lambda expression vector with *in vivo* excision properties. *Nucleic Acids Res.*, 16(15):7583-600.
- Sigler, P.B., Xu, Z., Rye, H.S., Burston, S.G., Fenton, W.A., Horwich, A.L. 1998. Structure and function in GroEL-mediated protein folding. *Annu. Rev. Biochem.*, 67:581-608.
- Sirevåg, R., Castenholz, R.W. 1997. Aspects on carbon metabolism in *Chloroflexus*. *Arch. Microbiol.*, 120:151-153.
- Strauss, G., Fuchs, G. 1993. Enzymes of a novel autotrophic CO₂ fixation pathway in the phototrophic bacterium *Chloroflexus aurantiacus*, the 3-hydroxypropionat cycle. *Eur. J. Biochem.*, 25:633-643.
- Studier, F.W., Dunn, J.J. 1983. Organisation and expression of bacteriophage T7 DNA. *Cold Spring Harbour Symp. Quant. Biol.* 47:999-1007.
- Studier, F.W., Moffat, B.A. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189:113-130.
- Synstad, B. 1988. Konstruksjon av genomisk bibliotek og søk etter genet for malat dehydrogenase i *Chloroflexus*. Hovedfagsoppgave i generell mikrobiologi ved avdelingen for Molekylær Cellebiologi, Biologisk Institutt, Universitetet i Oslo.
- Tagu, D. 1993. Identifying false positives during the first round of cDNA library screening. *Biotechniques*, 15(5):822.
- Taguchi, H., Yoshida, M. 1995. Chaperonin releases the substrate protein in a form with tendency to aggregate and ability to rebind to chaperonin. *FEBS Lett.*, 359:195-198.
- Takumi, T. 1997. Use of PCR for cDNA library screening. *Methods Mol. Biol.*, 67:339-44.
- Takumi, T., Lodish, H.F. 1994. Rapid cDNA cloning by PCR screening. *BioTechniques*, 17:443-444.

- Tanaka, N., Hiyama, T., Nakamoto, H. 1997. Cloning, characterisation and functional analysis of *groESL* operon from thermophilic cyanobacterium *Synechococcus vulcanus*. *Biochem. Biophys. Acta.*, 1343:335-348.
- Tooming, A.. 1998. Undersøkelse av termostabilitet hos malat dehydrogenase fra den termofile bakterien *Chloroflexus aurantiacus*. Hovedfagsoppgave i generell mikrobiologi ved Avdeling for Molekylær Cellebiologi, Biologisk Institutt, Universitetet i Oslo.
- Trent, J.D., Gabrielsen, M., Jensen, B., Neuhard, J., Olsen, J. 1994. Acquired thermotolerance and heat shock proteins in thermophiles from the tree phylogenetic domains. *J. Bact.*, 176:6148-6152.
- Vanberg, C. 2001. Isolering og karakterisering av genene for chaperoniner fra den termofile bakterien *Chloroflexus aurantiacus*. Hovedfagsoppgave i generell mikrobiologi ved avdeling for molekylær biologi, Biologisk Institutt, Universitetet i Oslo.
- Watson, K. 1990. Microbial stress proteins. *Avd. Microb. Physiol.*, 31:183-223.
- Weissman, J.S., Kashi, Y., Fenton, W.A., Horwich, A.L. 1994. GroEL-mediated protein folding proceeds by multiple rounds of binding and release of nonnative forms. *Cell*, 78:693-702.
- Woese, C.R., Kandler, O., Wheelis, M.L. 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria and Eucarya. *Proc. Natl. Acad. Sci. USA*, 87:4576-4579.
- Xu, Z., Horwich, A.L., Siegler, P.B. 1997. The crystal structure of the asymmetric GroEL-GroES-(ADP)₇ chaperonin complex. *Nature*, 388:741-750.
- Yu, L., Bloem, L.J. 1996. Use of polymerase chain reaction to screen phage libraries. *Methods Mol. Biol.*, 58:335-9.
- Yura, T., Nagai, H., Mori, H. 1993. Regulation of heat shock response in bacteria. *Annu. Rev. Microbiol.*, 47:321-350.
- Yura, T., Nakahigashi, K. 1999. Regulation of the heat shock response. *Curr. Opin. Microbiol.*, 2:153-158.

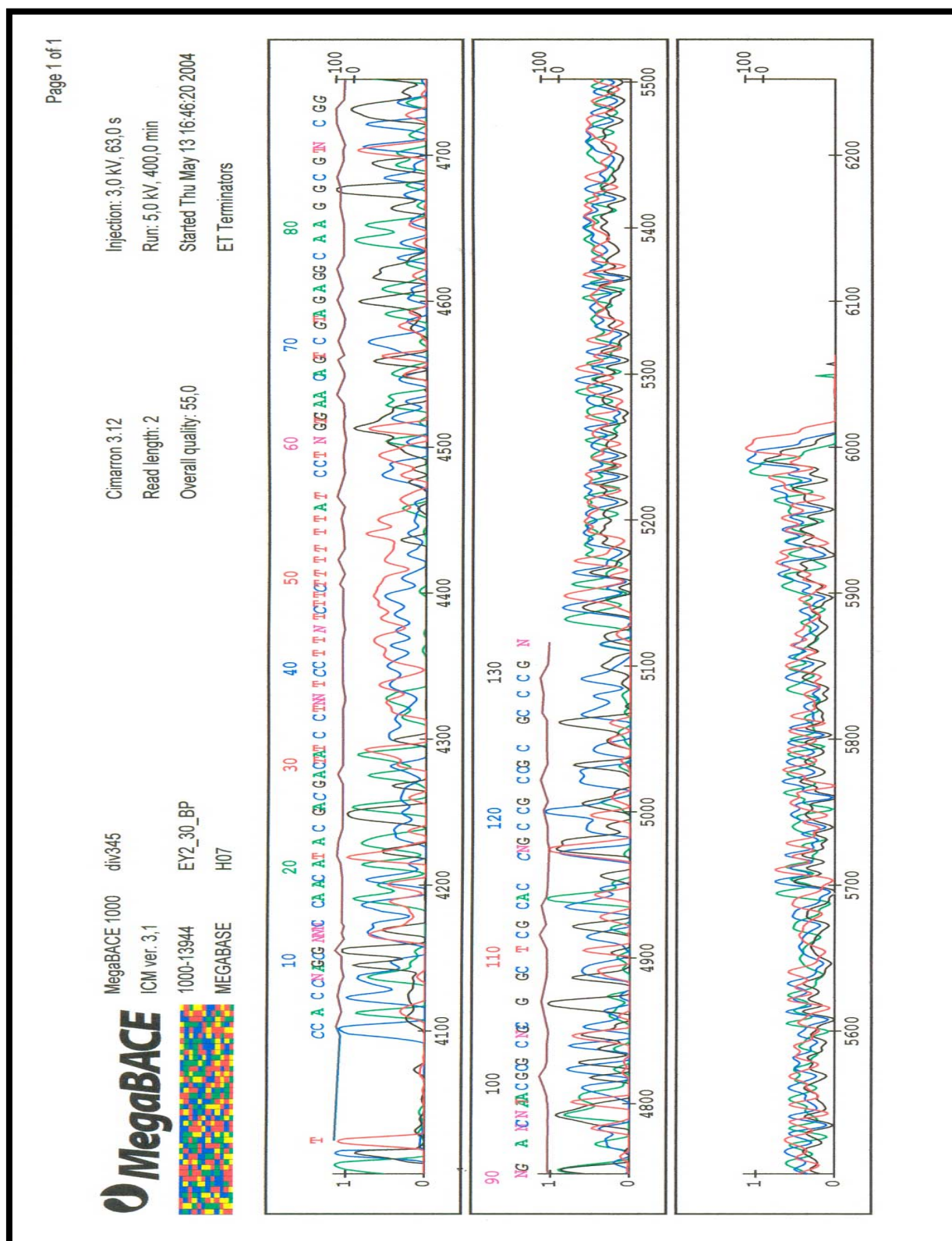
APPENDIKS



1. *groEL* 2. Chlo01_574 3. Chlo01_574 4. Chlo_724 5. Konsensussekvens



Kromatogrammet fra sekvensering med sekvenseringsprimer.



Kromatogrammet fra sekvensering av *groEL1*.

